

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	374	10.4	10.4	1958	12	US-10-028-946-4	Sequence 4, Appli
2	374	10.4	10.4	2054	12	US-10-028-946-2	Sequence 2, Appli
3	371	10.3	10.3	2055	9	US-10-017-216-4	Sequence 4, Appli
4	367	10.2	10.2	2053	9	US-10-017-216-2	Sequence 2, Appli
5	358.5	10.0	10.0	2125	10	US-09-919-172-29	Sequence 29, Appl
6	355	9.9	9.9	1597	9	US-10-017-216-6	Sequence 6, Appli
7	354.5	9.9	9.8	868	9	US-09-884-001-19	Sequence 19, Appl
8	353.5	9.8	9.7	677	10	US-09-745-763-168	Sequence 168, App
9	350	9.7	9.7	1641	9	US-10-017-216-5	Sequence 5, Appli
10	323	9.0	660	10	US-09-864-761-47959	Sequence 47959, A	
11	320	8.9	1045	10	US-09-815-242-10617	Sequence 10617, A	
12	317	8.8	2139	10	US-09-727-384-6	Sequence 6, Appli	
13	315.5	8.8	879	9	US-10-108-605-217	Sequence 217, App	
14	314	8.7	1179	10	US-09-815-242-13262	Sequence 13262, A	
15	309	8.6	1179	10	US-09-815-242-13608	Sequence 13608, A	
16	303	8.4	2310	9	US-09-931-436-120	Sequence 120, App	
17	303	8.4	2310	10	US-09-874-923-120	Sequence 120, App	
18	303	8.4	2568	10	US-09-866-108-3	Sequence 3, Appli	
19	297.5	8.3	909	10	US-09-925-299-988	Sequence 988, App	

Qy 232 EKLLYEIEISCASQVEKYKLDIAQLENLKKNDEILSKOSLEENIVILSKQVEDLN 291  
Db 601 RATECQHKLLKADQKQPEVGEYAKLEKINAQQLKIQELQEKLE-----KAVKAST 653  
Qy 292 VKCOLLE--KEKEDHVRNRNHNENLNAEQNLKQKFILEQOEHEKLOQKELQIDSLLQ 349  
Db 654 EATELLQNIQAKERAERELEKLONRDSESGIRKKLVEAEERRHSLNKKVRKLTMR 713  
Qy 350 EKELSSSLHQLKCSFOEEMVMEKNLFEERLKTDFELDKLOQKEQAERLVKOLEEAKS 409  
Db 714 ENRLKDDIQTK-----SQIQMAKDKILEEKKHREAOQVSAQHLEVHLKQ 758  
Qy 410 R----AEELKLLEELKLGKEAELE-----KSSAAHTQATILLQEK-----YDSMVOS 452  
Db 759 KEQHYEEKIKVLNDQIKDLADKYLENNMQRHEEAHEKGLISEQKAMINAMDSKIRS 818  
Qy 453 LEDVTAQESYKALPASE-----IEDLKLENSLQEKAAKAGNAEDVQHQ 498  
Db 819 LEQRIVSELSEANKLAANSLFTQNNKAQEEMISELRQOKFYLETQAGKLEAQNKRLEQ 878  
Qy 499 ILATESSNQYVRMLLDLOTKSALKETETKEITVSEFLQKITDLOLQKOOBEDF----- 552  
Db 879 LEXISHQDHSKDNRLLELETRLVREVSLEHEEQKLELKLQLTLEQLSLQERESQLTALQAA 938  
Qy 553 RKQLEDE--EGRKAEXENTTAELTEINKWRLLYEELYNKTKPFQQLDAFEVEKQALLNE 611  
Db 939 RAALESQRLQAKTELEETAAEEHQAHTAURDEI-----QRKFDALRNSCTVITD- 990  
Qy 612 HGAQEQLNKIRDSYAKLLGHQN--LKQIKHVVKLDENSLKSEVSKLRQLAKKKOS 669  
Db 991 ---LEEQLNQLTDEDAE-LNNQNFYLSKOLDEASGANDEIVOLRSEVDHLRREITEREMQ 1046  
Qy 670 ETKLOEEL 677  
Db 1047 LTSQKQTM 1054  
RESULT 2  
US-10-028-946-2  
; Sequence 2, Application US/10028946  
; Patent No. US20020123622A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Miranda, Maricar  
; APPLICANT: Friddle, Carl Johan  
; TITLE OF INVENTION: NO. US20020123622A1el Human Kinases and Polynucleotides Encoding  
; FILE REFERENCE: LEX-0289-USA  
; CURRENT APPLICATION NUMBER: US/10/028,946  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: US 60/258,335  
; PRIOR FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 2054  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-028-946-2  
Query Match 10.4%; Score 374; DB 12; Length 2054;  
Best Local Similarity 21.2%; Pred. No. 3e-09;  
Matches 154; Conservative 155; Mismatches 297; Indels 122; Gaps 20;

```

Qy      11  FNDP-----SGCAPSGAYDVKTLEVLKGPVFSQSRFKQKESQKNLVNDKDTTL 62
      388  FDEPKNSWSSPQLSPGSGFGBELPFV--GPSYKALGILGRSES-----VVSGLDS 440
Db
Qy      63  PASARKYSSSEKSESORNDKLTLEKEIRVLQERGAQDPRIODLTELEKMEARLNA 122
      441  PA-----KTSMEKRLTKSELQDSQDKCHKMEQMTFLHRRVSEVAVLQSKEVELKA 495
Db
Qy      123  ALREKTLISANNATLEKOLIELTRTNELLKSKFSGNGKNLRIILSLEMLKLRNRETKM 182
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :

```

Db 496 SETQRSLLEQDLATYITTECSLSKLSLEQARMEVSEQDDK-----ALQLLHDIREQSRKL 549  
 Qy 183 RGMMAKQGBMEMKLQVQTSRSLSESGKTAQLEKGLVS-----IEKEFIDEKSET 231  
 Db 550 Q-----BIKEQEYQAQVVEWRMLMNQLEEDLVSNRRSDLYESELRESRLAAEEFK 600  
 Qy 232 EKLLEYIEEISCADQVEKYKLDIAQLAELENKEKNDIEILSKQSLERNIVILSKOVEDLN 291  
 Db 601 RKATECQHKLLKAKDQGPVEGYAKLEKINAEQOLKIQELQEKLE-----KAVKAST 653  
 Qy 292 VKCOLLE--KEKEDHVNRRHENNLNNAEQNLKQKFTLEQOEHEKLOOKELOIDSLLOQ 349  
 Db 654 EATELLQINQRAERAERELEKLNQREDSSGIRKKLVAEERHSLSENKVKRLTWMRR 713  
 Qy 350 EKELSSSLHQKLCFSQEBEMVKEKNLFEEELKOTDELDKLOQKEQAEERLVKQLEEEAKS 409  
 Db 714 ENRLKDDIQTk-----SIOIQOMADKLTILEEKEHREQAQVSAQHLEVLHKQ 758  
 Qy 410 R----AEELKLEBKLEKGAELE-----KSSAAHTQATLLLOEK-----YDSMVQS 452  
 Db 759 KEQHYEEKIKVLNDQIKKDLADKLTENMMQRHEEAHEKGKILSEQKAMINAMDSKIRS 818  
 Qy 453 LEDVTAQFESYKALTASE-----IDELKLENSLQEKAKAKGNAEDVQHQ 498  
 Db 819 LEQRIVELSEANKLAANSLSFTORNMKQAQEMISELRQQRKYLETQAGKLEAQNRLKEEQ 878  
 Qy 499 ILATESSNOEYVRMLDLDTKSALKETEIKETIVSYFLQKITDNLQNLKQOEDE-----552  
 Db 879 LEKISHQDSKNRLELETBLREVSLEHEEQKLELKRQLTELOLSQERESQALTALQAA 938  
 Qy 553 RKQLEDE--EGRKARKENTTAELTEINKWRLLYBELYNKTKPFQQLDQDAFEVEKQALLNE 611  
 Db 939 RAALESQLRQAQKTELETTAEAEIEQALTAHRDEI-----QRKFDALRNSCTVTD- 990  
 Qy 612 HGAQAEQLNKTDRSYAKLLGHQ--LKQKIKHVYKLDENSQKSEVSKLRCQLAKKKQS 669  
 Db 991 ---LEEQLNQLTENDAE--LNNQNFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQ 1040  
 Qy 670 ETKLQEEL 677  
 Db 1047 LTSOKQTM 1054  
 RESULT 3  
 US-10-017-216-4  
 ; Sequence 4, Application US/10017216  
 ; Patent No. US20020160483A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KAPLELLER-LIBERMANN, Rosana  
 ; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type P  
 ; TITLE OF INVENTION: Kinase and Uses Thereof  
 ; FILE REFERENCE: 10147-57U1  
 ; CURRENT APPLICATION NUMBER: US/10/017,216  
 ; CURRENT FILING DATE: 2001-10-23  
 ; PRIOR APPLICATION NUMBER: US 60/242,429  
 ; PRIOR FILING DATE: 2000-10-23  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 2055  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-017-216-4

```

Query Match      10.3%   Score 371;   DB 9;   Length 2055;
Best Local Similarity 21.7%;   Pred. No. 4e-09;
Matches 159;   Conservative 151;   Mismatches 291;   Indels 132;   Gaps 23;

QY 11 FNDPSG-----CAPS-PGAYDVKTLLEVLKGPVSFQKSRFQKQKESKONLVNDKDTTL 62
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 387 FDEPEKNSWAFILCVPAEPLAFSGEELPFV--GFYSKALGYLGRSES-----VSSLDS 439

QY 63 PASARKVKSSSEK-----KESQNKDKDLKLEIKRVLVLLQERGAQDRRIIDLETSEKME 117

```

```

Db 440 PA-----KVSMKKLLIKSKELQSDQKCHKEQEMTRL-----HRRVSEVEAVLSQKE 489
QY 118 ARNLAAREKTSLSANNATLEKQLIELTRTNELLKSFSENGQKNLRILSLMLKRNK 177
Db 490 VELKASETQSLLEQDIATYITECSSUKRSLEQARMEVSQEDDK-----ALQLLHDIRE 543
QY 178 RETKMRGMAKQEGEMKQVQTSRLESQKIAQLEKGLVS-----IEKEKID 226
Db 544 QSRKIQ-----EIKEQYQAOVEEMRLMNMQLLEEDLVARRRSDLYESELRESRIA 594
QY 227 EKSTTEKLYIEBISCASQVQYKIDIAOLEENLEKNDIEILSKOSLEENIVILSKQ 286
Db 595 AEEFKRANECQHKLMKADQKQKPEVGEYSKLEKINAEQOLKIOLEKLE-----KA 647
QY 287 VEDLVKCOLLE--KEKEDHVRNREHNENLNAEMNLKOKFLEQOEHEKLOQKELQID 344
Db 648 VKASTEATELLQNRQAKRAERELEKLNHRNEDSSGIRKKLVAEERRHSLENKVRLE 707
QY 345 SLIQOEKELSSSHOKLCSFQEEEMVKEKNLFEELKOTLDLQKQKEQAEALRVKOLE 404
Db 708 TMRERENLKDIOTK-----SQIOQMAKILEEKEKHREAOVSQAHLE 752
QY 405 EAKSR-----AEELKLEELKQKEALEK-----SSAAHTQATLLLOEK-----YD 447
Db 753 VHLKQEKQHYEEKIKVLNDQIKKDLADKESLNNMQRHEEAEHKGKILSEQKAMINAMD 812
QY 448 SMVQSLEDVTAQESYKALTASE-----IEDKLNSLSLOEKAAGNAE 493
Db 813 SKIRSLQRIVESEANKLAANSSLFQTNMKQAEEMISELROQKFYLETQAGKLEAQRN 872
QY 494 DVHQIILATESNOEYVRMLLDLQTKSALKETEITVTSFLOKITDLOLQKQOEDF- 552
Db 873 KLEBQEKISHODSKRSRLELETRLEVSLEHEEQKLEKRLQTLQSLQRESQLT 932
QY 553 -----RKOLEDE--EGRKAQKENTAELETEINKWRLLYEELYNKTKPFQLOLDAFEVEKQ 606
Db 933 ALQAAALAESQLROAKTELETTAEAEETQALTAHRDEI-----QRKFDALRNSCT 985
QY 607 ALLNEHGAEOELNKIRDSYAKLLGHON--LKOKIKHVVKLKDENSOLKSEVSKLRQOLA 664
Db 986 VITD-----LEEQLNLTEDNAE--LNNQNFYLSKOLDEASGANDEIVOLRSEVDHLRREIT 1040
QY 665 KKQOSETKLOEEL 677
Db 1041 ERENQLTSQKQTM 1053

```

```

RESULT 4
US-10-017-216-2
; Sequence 2, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Pro
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-216-2

```

```

Query Match 10.2%; Score 367; DB 9; Length 2053;
Best local similarity 21.0%; Pred. No. 6e-09;
Matches 153; Conservative 151; Mismatches 287; Indels 136; Gaps 22;

```

```

QY 11 FNDP-----SGCAPSPGAYDVKTLEVLKGPVSFQKSRQKQKESQKNLNDKDDTTL 62
Db 388 FDEPKNSWSSVSPQLSPSCFSGEELPFV--GFSYKALGILGRSES-----VVSGLDS 440
QY 63 PASARKYKSESSESQKNDKDLKILEKEIRVLQOERGAQRRRIODLETELEKMEARUNA 122
Db 441 PA-----KTSMEKKLLIKSKELQSDQKCHKEQEMTRLHRRVSEVEAVLSQKEVELKA 495
QY 123 ALREKTSLSANNATLEKOLITELTRTNELLKSFSENGQKNLRILSLMLKRNKRETKM 182
Db 496 SETORSLEQDLATYITECSSUKRSLEQARMEVSQEDDK-----ALQLLHDIRESRKL 549
QY 183 RGMMAKQEGEMKQVQTSRLESQKIAQLEKGLVSTIEKEKIDKESKTEKLEYIERIS 242
Db 550 Q-----EIKEQYQAOVEEMRLMNMQLLEEDLVARRRSDLYESE-----LRESR 593
QY 243 CASQVQYKIDIAOLEENLEKND-----EILSKOSLEENIVILSKQVEDL 290
Db 594 LAA---EERKRAKATECQHKLLKADQKQKPEVGEYAKLEKINAEQOL-----KIOEL 641
QY 291 NVKCOLLEKEDHVRNREHNENLNAEMNLKOKFLEQOEHEKLOQKELQIDSLQOE 350
Db 642 QEK---LEKAARAEARELEKLNHRNEDSSGIRKKLVAEERRHSLENKVRLETMERRE 698
QY 351 KESSSHOKLCSFQEEEMVKEKNLFEELKOTLDLQKQKEQAEALRVKQLEEEAKSR 410
Db 699 NRLKDDIOTK-----SQIOQMAKILEEKEKHREAOVSQAHLEVHLKQK 743
QY 411 -----AEELKLEELKQKEALE-----KSSAAHTQATLLLOEK-----YDSMVOSL 453
Db 744 EQHYEEKIKVLNDQIKKDLADKESLNNMQRHEEAEHKGKILSEQKAMINAMDSKIRSL 803
QY 454 EDVTAQESYKALTASE-----IEDKLNSLSLOEKAAGNAEDVOHQI 499
Db 804 EQRIVELSEANKLAANSSLFQTNMKQAEEMISELROQKFYLETQAGKLEAQRNKLLEQL 863
QY 500 LATESNOEYVRMLLDLQTKSALKETEITVTSFLOKITDLOLQKQOEDF-----R 553
Db 864 EKISHODSKRSRLELETRLEVSLEHEEQKLEKRLQTLQSLQRESQLTALQAA 923
QY 554 KOLEDE--EGRKAQKENTAELETEINKWRLLYEELYNKTKPFQLOLDAFEVEKQALLNEH 612
Db 924 AALESOLROAKTELETTAEAEETQALTAHRDEI-----QRKFDALRNSCTVITD-- 974
QY 613 GAAEOELNKIRDSYAKLLGHON--LKOKIKHVVKLKDENSOLKSEVSKLRQOLAKKQSE 670
Db 975 --LEEQLNLTEDNAE--LNNQNFYLSKOLDEASGANDEIVOLRSEVDHLRREITEREMQL 1031
QY 671 TKLOEEL 677
Db 1032 TSQKQTM 1038

```

```

RESULT 5
US-09-919-172-29
; Sequence 29, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 2125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 3774181CD1
US-09-1919-172-29

Query Match      10.0%; Score 358.5; DB 10; Length 2125;
Best Local Similarity 21.0%; Pred. No. 1.4e-08;
Matches 179; Conservative 157; Mismatches 268; Indels 249; Gaps 37;

Qy 26 VRTLVLPKGPVSQSFQKQKESQKLNVD-----KDTTLP-----ASA 66
   | : : : : : | | | | | : : : : : | | : : : : : |
Db 764 LKLIQKMEKDIAFQKQAEKQKE-KQIELEARRKITEIQTCEENALPVPITQATSC 822
   | : : : : : | | | | | : : : : : | | : : : : : |
Qy 67 RKVKSSEKSKQNKDKLKIKEIRVLLQERGAQDRRIODLETELEKMEARLNAALRE 126
   | : : : : : | | | | | : : : : : | | : : : : : |
Db 823 RAVTGLQOEHDQK-----AELKQVDELTAANK---AEQDMRELYELNALQLE 871
   | : : : : : | | | | | : : : : : | | : : : : : |
Qy 127 KTSLSANNATLEKQLIETRTNELLKSF-----SENGNQNLRLSLLEMLKLNKRETK 181
   | | | : : : | : | : | : : : | : | : | : : : |
Db 872 KTSSEKARLLKLDETNNTLRLCLLEKRDQAEKGYSQLRGRLQNTTGGKAEAA 931
   | : : : : : | | | | | : : : : : | | : : : : : |
Qy 182 MRGMAKQEGMEMK-----LQVTORESOGKIAQLEGKLVSTKEKIDE-KSETKLL 235
   | : : : : : | | | | | : : : : : | | : : : : : |
Db 932 M-----QASDLKKIKKNYOLESLNHEKQKQKQ-----EVDRIIRAHAFAEKNI 977
   | : : : : : | | | | | : : : : : | | : : : : : |
Qy 236 EYI-EISCASDOVEKYKLDIAQ-----LEENLKEKNDIL-SLKQSLLEN--IVILSKQ 286
   | : : : : : | | | | | : : : : : | | : : : : : |
Db 978 QHLNLSQIHSFRDEKELERLQICQKSDHLKEQFEKSHQOLLQNTKAEKENDKIORLNEE 1037
   | : : : : : | | | | | : : : : : | | : : : : : |
Qy 287 VEDLVNVCOLLEKEDHVNRRHN-----ENLNAEMQNLKQKFILEQOE-HEKLOQKLEQ 342
   | : | : : : | : | : | : : : | : | : : : : |
Db 1038 LEKSNCEAEMLKQKVEELTRONNETKLMQMQRIQAESENI-----VLEKQTIQORCEALKIQ 1093
   | : | : : : | : | : | : : : | : | : : : : |
Qy 343 IDSLLQKEKELSSSLH-----QKLCFQEMVMVEKNLFEELKQTLDELDKLOOK 392
   | : | : : : | : | : | : : : | : | : : : : |
Db 1094 ADGFKDLRSTNHLHKTQTEQDFQRIKCLEEDLAKSONLV-SEFKQKCDQOQNIQ- 1151
   | : | : : : | : | : | : : : | : | : : : : |
Qy 393 EQAERLVKQLEEAASRAELKLEELKQKEAELEKSS-----AAHTQAT 439
   | : | : : : | : | : | : : : | : | : : : : |
Db 1152 ---NTRKKEVRLNAELNASKEERKRGQKQVQLQAQVQELNNRLKQVDELHLKLTTEEOMT 1209
   | : | : : : | : | : | : : : | : | : : : : |
Qy 440 ---LLQEKYDSMVQSLVEDVTAQFESY---KALTASIEDLKLSENSLOEKAAGKNA 492
   | : | : : : | : | : | : : : | : | : : : : |
Db 1210 HRKMVLFQESGKFKQSAEEFKKMEKLMESKVITENDISGIRLDFVSLQOENSRAQENA 1269
   | : | : : : | : | : | : : : | : | : : : : |
Qy 493 EDVQHQILATSSNOEYVRMLLDLQTSALKATEIKETITVSFLOKITDLQNLKQOEEDF 552
   | : | : : : | : | : | : : : | : | : : : : |
Db 1270 K-----LCETNIKE-----LERQLQOYREOM 1290
   | : | : : : | : | : | : : : | : | : : : : |
Qy 553 R-----KQLEDE-EGRAKAEKNTAELTEINKWRLLYEELYNKTKPFQLOL 598
   | : | : : : | : | : | : : : | : | : : : : |
Db 1291 OQGQHMEANHYKQCKLEDELIAQKREVENLKQMDQOI-----KEHEHQLVL 1338
   | : | : : : | : | : | : : : | : | : : : : |
Qy 599 DAFEVEKQALLN-----EHGAQAEQLNKIRDSYAKLLGH----- 632
   | : | : : : | : | : | : : : | : | : : : : |
Db 1339 LOCEIQKSTAKDCTFKPDFEMTVKECHSGELSSRN-----TGHLPHTPRSPLLR 1389
   | : | : : : | : | : | : : : | : | : : : : |
Qy 633 ---ONLKQIKIH-----VVKLKDENSQKSEVSKRCOLAKKQSETKLQ-----EE 676
   | : | : : : | : | : | : : : | : | : : : : |
Db 1390 WTQEPQPLEEKWOHRVVEIQIPEKVFQPPGAPLEKESQ-QCYSEYFSQSTELQITFDE 1448
   | : | : : : | : | : | : : : | : | : : : : |
Qy 677 LNKVLGKIHFD-----PSKAFHESKENFALK--TPLEKGNVNCY----- 714
   | : | : : : | : | : | : : : | : | : : : : |
Db 1449 TNPIRFLSEIKIRDOALNNSRPVRYODNACEMELVKVLPLEIAKNQYDMHTEVTTL 1508
   | : | : : : | : | : | : : : | : | : : : : |
Qy 715 ---RAPMECQESW 724
   | : | : : : | : | : | : : : | : | : : : : |
Db 1509 KOENKPNVPSAEW 1521
   | : | : : : | : | : | : : : | : | : : : : |

Query Match      9.9%; Score 355; DB 9; Length 1597;
Best Local Similarity 21.4%; Pred. No. 1.5e-08;
Matches 137; Conservative 139; Mismatches 262; Indels 102; Gaps 16;

Qy 91 EIRVLLQERGAQDRRIODLETELEKMEARLNAALREKTSLSANNATLEKQLIETRTNREL 150
   | : : : : : | | | | | : : : : : | | : : : : : |
Db 5 EAMMEQEMTLRHRRVSEVEAVLSQKEVELKASQTQRSLLQODLATYITESSLKRSLEQ 64
   | : : : : : | | | | | : : : : : | | : : : : : |
Qy 151 LKSFSENGNQNLRLSLLEMLKLNKRETKMRGMMAKQEGMEMKLVQVTORESLESOQKI 210
   | : : : : : | | | | | : : : : : | | : : : : : |
Db 65 ARMEVSOEDDK-----ALQLLHDIREQSRKLQ-----EIKEQEQYQAOVEEMRLMM 109
   | : : : : : | | | | | : : : : : | | : : : : : |
Qy 211 AOLLEGKLV-----IEKEKIDEKSETEKLELYIEISCASDOVEKYKLDIAOLE 259
   | | | | | : : : : : | : : : : : | : : : : : |
Db 110 NQLEEDLVARRRSDLYESLRESRLAAEFKANKQCQHKLMKAKQDGKPEVGEYSKLE 169
   | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 260 ENLKEKNDILSKLSLEENIVILSKOVEDLVNVCOLLE---KEKEDHVNRRHNENLNA 317
   | : : : : : | : : : : : | : : : : : | : : : : : |
Db 170 KINAEOQLKIOELQEKLE-----KAVKASTEATELLONIROKARERLEKLNRED 222
   | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 318 EMQNLKQFILEQHEKQKELQIDSLQOEKELSSSLHOKLCSFOEEMVKEKNLFEE 377
   | : | : : : | : | : : : | : | : : : | : | : : : |
Db 223 SSEGIKKKLVAEERRHSLNENKVRKLETMERRENRLKDDIOTK-----SE 267
   | : | : : : | : | : : : | : | : : : | : | : : : |
Qy 378 ELKQTLDELQKQKEQAERLVKQLEEAASR-----AEELKLEELKQKEAELEK--- 430
   | : | : : : | : | : : : | : | : : : | : | : : : |
Db 268 QIQQMAQKILELEKHREAOVSAQHLVHLKQKEQHYEEKIKVLDNQIKKLADKESLEN 327
   | : | : : : | : | : : : | : | : : : | : | : : : |
Qy 431 ---SSAAHQATLLQEK-----YDSMVQSLVEDVTAQFESYKALTASE----- 470
   | : | : : : | : | : : : | : | : : : | : | : : : |
Db 328 MQQRHEEAHEKGLISEOKAMINAMDSKIRSLQRIVELSEANKLAANSSLTQRMKA 387
   | : | : : : | : | : : : | : | : : : | : | : : : |
Qy 471 ---IEDLKLENSLQEKAAKAGNAEDVQHQILATESSNOEYVRMLLDLQTSALKRETE 526
   | : | : : : | : | : : : | : | : : : | : | : : : |
Db 388 QEEMISELRQOKFVLETQAGKLEAQNRLKEQLEKISHQDHSKSRLLLETRLSVSLE 447
   | : | : : : | : | : : : | : | : : : | : | : : : |
Qy 527 IKETIVSFLOKITDLQNLKQOEEDF-----RKQLEDE-EGRAKAEKNTAELTEELINK 579
   | : | : : : | : | : : : | : | : : : | : | : : : |
Db 448 HEEOKLEKRLQTLQELQSLQERESQLTALQARAALLESQLRQAQKTELEETAAEEBEOA 507
   | : | : : : | : | : : : | : | : : : | : | : : : |
Qy 580 WRLLYEELYNKTKPFQLOLDAFEVEKQALLNEHGAQAEQLNKIRDSYAKLLGHON--LKQ 637
   | : | : : : | : | : : : | : | : : : | : | : : : |
Db 508 LTAHRDEI-----QRKFDALRNSCTVITD-----LEEQLNQLTEDNAE-LNNQNFYLSK 555
   | : | : : : | : | : : : | : | : : : | : | : : : |
Qy 638 KIKHVVKLDNSQKSEVSKRCOLAKKQSETKLQEEEL 677
   | : | : : : | : | : : : | : | : : : | : | : : : |
Db 556 QLDEASGANDEIVQLRSEVDHLRREITERENMLTSQKQTM 595
   | : | : : : | : | : : : | : | : : : | : | : : : |

RESULT 7
US-09-884-001-19
; Sequence 19, Application US/09884001
; Publication No. US20020182656A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; Peschon, Jacques J.
```



APPLICANT: Sims, John E.  
APPLICANT: Virca, G. Duke  
APPLICANT: Willis, Cynthia R.  
TITLE OF INVENTION: Methods for Regulating Vascularization Using GEF  
TITLE OF INVENTION: Containing NEK-Like Kinase (GNK)  
FILE REFERENCE: Immunex GNK/sgnk pct  
CURRENT APPLICATION NUMBER: US/09/884,001  
CURRENT FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 60/113,003  
PRIOR FILING DATE: 1998-12-18  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn ver. 2.0  
SEQ ID NO 19  
LENGTH: 868  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-884-001-19

Query Match 9.9%; Score 354.5; DB 9; Length 868;  
Best Local Similarity 22.0%; Pred. No. 9.4e-09;  
Matches 190; Conservative 150; Mismatches 293; Indels 229; Gaps 31;

Qy 25 DVKTLEVLKGPVQSKORFQKQESKONLVNDKDTLPASARKVKSESKESQKNDK 84  
Db 21 DLRIQOEKEIKKLSRQGEAA-----TTQLEQLHQAQKQEEVLARAVQKE 72  
Qy 85 LKILEK---EIRVLLOERGAQD--RRIQDLELEKMEARLNAALREKTSLSANNATLEK 139  
Db 73 ALVREKAALEVLQNAVERDQDLAQLGSLSSAKELLESLFEAQQQSVIVTKGQLEV 132  
Qy 140 QLIETRTNELLKSKFSENGKQNLRIYLSLEMLKRNKRETKMRG-----MMAKQEG--- 191  
Db 133 QIQVTQAKEVI-----QGEVRLKLELDTERSQAQERDAARAAQLAQAEQEGTKA 183  
Qy 192 -----MEMKLOVQORSLEESQ 208  
Db 184 LEQKAAHEKVNOLREKWEKERSWHOQELAKALESREKMELEMLRLEKQOTTEAIG- 242  
Qy 209 KIAQEGKLVISIEKIDKSETEK-----LLEYTEISCSADQVEKYKLDI----- 255  
Db 243 --AQREERTQASALCOMQLETEKERSVILLETLLQTKQLADASQQLERLRQDMKVOKL 300  
Qy 256 -----AQLEENLEKNDLSLKQ---SLEENIVILSKQVEDLNVKC- 294  
Db 301 KEQETTGILQLOEAQRELEKAAQRHDDLAALQEESSLLQDKMDLQKQVEDLKSQLV 360  
Qy 295 -----QLLEKEKEDHVNREHN---ENLNAEQNL-----TKQFILEQOEHEKLOQK 339  
Db 361 AQDQSRLVEQVEQKLEKRETOEYNNRIQKELERERKASLTLSMERKFORLLVLQEAADSI 420  
Qy 340 ELQIDSLQOEKELSSSLHQKLCFOREM---VREKNL-PEEELKOTLDLDELKLOQKEE 394  
Db 421 EL---SALRQDMQEAQG-EQKELSAQWELLQKVEKEADFLAQEAQLLEELASHITEQ 476  
Qy 395 QAERLVKOLEBEAKSRABELKLEKLGKBAELEKSSAAHTQATL-----LLAQEKYDSM 449  
Db 477 QLRASLWAEAKAAQQLRLRSTESQLEALAAEQPGQAQQAQALASLYSALQQAQALGSV 536  
Qy 450 VOSLEDVTAQTES-----YKALPASEI-----EDL----- 474  
Db 537 CESRPELGGGDSAPSVNGLPEPDONGARSLFKRGPPLLTALSAEAVASALHKLHODLWK 596  
Qy 475 -----KLENSISQEAAGKNAE--DVQHQILATYESSNOEYVRLMLDLOTKSA 521  
Db 597 QTRDVLQVOKLEERLTDTEAKSKSVHTELQDLQORQL-----SONQEE-----KSKWE 645  
Qy 522 LKTEIKEITVSFQKIDTDLQNLQKQEDFRKQLEDEEGRAKENTTA----- 571  
Db 646 GKQNSLESELMELHETMASLQSRRLRAELQRMQAQGERELLQAAKENLTAQVEHLQAAVV 705  
Qy 572 ELTEINKRWRLLYELYNKTKPFQLOLDAFEVEKQALLNEHGAQAEQNLKTRDSYAKLLG 631  
Db 706 EARAQASAAAGILEEDL--RTARSALKLNKEEVESE---RERAQALQOEQELKVAQKALQ 760

Qy 632 H-----ONLKQIKHVVYKLDKENSOLKS--EVSKLRCQL-----AKKQOSETKIQ-----E 675  
Db 761 ENLALLTQTLAEREVEVETLRGQIOELEKQREMQRAALELLSLDLKKRNQEVLDLQOEIQ 820  
Qy 676 ELNKVLGIKHFDPDSKAFHESK 697  
Db 821 ELEKCRSVLEHLPLMAVQEREQK 842  
RESULT 8  
US-09-745-763-168  
Sequence 168, Application US/09745763  
Patent No. US20020065394A1  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
LaVallie, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 168:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 677 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 168:  
US-09-745-763-168

Query Match 9.8%; Score 353.5; DB 10; Length 677;  
Best Local Similarity 25.0%; Pred. No. 8.4e-09;  
Matches 175; Conservative 126; Mismatches 242; Indels 157; Gaps 33;

Qy 69 VKSESKEKESQK-----NDRDLKILEK-----EIRVLQERGAQDRRIQDLETEL--- 113  
Db 8 LKVSDEMKKSHDAITDDLNRKLLDVTQYTKKLEMEKLLLENDLSLSDVSRLFTVFVPP 67  
Qy 114 EKMEARLNAALREKTSLSANNATLEKQILTELTR-----FNEILKSKFSENGKQNLRIISL 169  
Db 68 EKHE-----KEITALSKSNIVELKQLSELKKKCGEQEKIHALTSENTNLK--KMMSN 118  
Qy 170 ELMKLRNKRKTKM--RGMMAK-----QEGMEM--KLOVQORSLEESQK 209  
Db 119 QVVPVKTHEEVKTLNDTLAKTNRELLDVKKKFEDINQEFVKIKDKNEILKRNLENTQ 178

QY 210 I-----AQEGKLVISIEK--EKIDKSETEKLLLEY-----IBEISCASQDQVEKYKLDIAQ 257  
 Db 179 IKAEXISLAHEAKKSSLSQSMRKVOD-SNAELIANYRGQEEIVTLHAEIKRAQKKEJDT 237  
 QY 258 LEENLKEKDEILSKQSLEENIVILSKQVEDLNKVCOLLEKEK-----EDHVNNRHNN 312  
 Db 238 IQEIKVKYAPVIFSEE-CERFKATEKELKD-----QLEQTKYKYSVSEEEVKNKQBN 291  
 QY 313 ENLNARM-----ONLQKFTLEQOEHEKLOQKELQIDSLQOEKELSSSLHOKLCSFQEBM 368  
 Db 292 DKLKKEIFLQDLDRKDTVLIEKSHMERALSRRKTDENKQKLDLS---QKYTEVKN-- 345  
 QY 369 VREKNLFEELKOTDEL--DKLQOE-----EQAEVLVKQLEEEAKSRAEELK-----415  
 Db 346 VREK-LVEENAKOTSILAVQNLQKHVPLEQVEALKSLNGTIENTLKEELKSMQRCVE 404  
 QY 416 -----LLE-----EKLKGEA-----ELEKSSAAHTQATLL 441  
 Db 405 KQQTQVTKLHQLLENQKNSVPLAEHLQIKFAKEVGIKASLREKEEESQNMKEEYSK 464  
 QY 442 LQEKDSVMYQSLDV-----TAQFESYKALTASEIEDLKLSSLOEKAAGKAGNAEDVQH 497  
 Db 465 LQSEVQNTQALKKLETRVVDLSYKATKS-----DLEQISSLNEKLANLNKRYEEVCE 520  
 QY 498 QILATE-----SSNQEVYRMLDLQTSALKETEIKETVTSFLQKITDQNLKQOEEDFR 553  
 Db 521 EVLHAKKEISAKDE--KELHFSIEQEIQDQ--KERCDKSLTTITELQRRIOES---A 572  
 QY 554 KQLEDEEGKAKENTTAELTEINKRWLLYEELNKTYPFQLOLDAFEVEKQALLNEHG 613  
 Db 573 KQIEAKDNKITELNDVERLQKALGLSOLTYTSGNPTKRSQDITLQHOVKSLEQQLA 632  
 QY 614 AAOEOLNKRDSYAKLL-----GH--ONLKQIKHVVKLK 646  
 Db 633 DADRQHQEVIATRYTHLLSAAOGHMDVEDVQEAALLQIQMR 672

RESULT 9  
 US-10-017-216-5  
 ; Sequence 5, Application US/10017216  
 ; Patent No. US20020160483A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KAPPELLER-LIBERMANN, Rosana  
 ; TITLE OF INVENTION: 13245, A No. US20020160483A1 Human Myotonic Dystrophy Type Pro  
 ; FILE REFERENCE: 10147-5701  
 ; CURRENT APPLICATION NUMBER: US/10/017,216  
 ; PRIOR FILING DATE: 2001-10-23  
 ; PRIOR APPLICATION NUMBER: US 60/242,429  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 1641  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-017-216-5

Query Match 9.7%; Score 350; DB 9; Length 1641;  
 Best Local Similarity 22.13%; Pred. No. 2.6e-08;  
 Matches 158; Conservative 140; Mismatches 285; Indels 126; Gaps 23;

QY 11 FNDP-----SGCAPSPGADVTKTLVKGVPVFSQKRSQKQESQNLNVKDFTTL 62  
 Db 15 FDEPKNSWSSVQCLSPSGSGELPFPV--GFYSYKALGVLGRSES-----VVSIDS 67  
 QY 63 PASARKVSKESK-----KESQNKDKDLKILEKEIRVLQERGAQDRRIODLETELEKME 117  
 Db 68 PA---KVSSMEKKLITKSKELQDSQDKCKHMQEOTRL-----HRRVSEAVLSQKE 117  
 QY 118 AFLNALREKTSLSANNATLEKQLELTFTNELLKSKSENGNQNLRLISLEMLKLNK 177

Db 118 VELKASQTORSLEQLDATTYITECSSLKRSLSEARMEVSEQEDDK-----ALQLLHDIRE 171  
 QY 178 RETKMRGMMAQEGMEMKLOVTSRLESOGKIAOLEGKLVSTIEKEKIDKSETEKLLLEY 237  
 Db 172 QSRKIQ-----EIKEQEYQAQVEEMRLMMNQLQLEEDLVSAARRSDLYESE-----215  
 QY 238 IBEISCASQDQVEKYKLDIAQLEENLKEKNDEILSKQSLEENIVILSKQVEDLNKVCOLL 297  
 Db 216 -----LRESLAAAEFKRANECQHKLMKADQCKPVEGVYSK-LEKINAEQOLK 264  
 QY 298 EKEKEDHVNRRNHNENLNAEMQNLKQFLEQOEHEKLOQKELQIDSLQOEKELSSSL 357  
 Db 265 IQELQEKLEKAVKASTATELQNTQAKERAERELEKLNRE-----DSSEGI 313  
 QY 358 HOKLCSFOEEMVKEKNLFEELKQTDDEL--KLOQKEQAEVLVKQLEEEAKSRAEELK 416  
 Db 314 KKLVEAE--LEEKH---REAQVSAQHLVHLKQEHYEEKIKVLDNQIKKDLAKRES 368  
 QY 417 LEEKLKGAEAELEKSSAAHTQATLLQEK-----YDSMVQSLSDVTAQFESYKALTASE- 470  
 Db 369 LENMQRHEE-----AHEKGKILSEQKAMINAMDSKIRSLERQIVELSEANKLAANSS 422  
 QY 471 -----IEDLKLESSLQEKAAKAGKNAEDVQHQILATLESSNOEYVRMLDLQ 517  
 Db 423 LFTORMKAQEMISELQKQFPYLETQAGKLEAQNRLKEEQLEKISHQDHSKSLLELE 482  
 QY 518 TKSALKETEIKETVTSFLQKITDQNLKQOEEDF-----RKQLEDE--EGRKAKENTT 570  
 Db 483 TFLREVSLSEHEQKLELKRQLTELQLSQERESQLTALQAAARAALESQLRQAKTELEETT 542  
 QY 571 AELTEINKRWLLYEELNKTYPFQLOLDAFEVEKQALLNEHGAQOEQLNKRDSYAKLL 630  
 Db 543 AEABEEIQAHTARDEI-----QRKFDALRNSCTVITD-----LEEQLNQLTDEANB-L 590  
 QY 631 GHON--LQKIKHVVKLKDENSQKSEVSKRLCOLAKKKQSETKLQEEEL 677  
 Db 591 NNQNFYLSQKDEASGANDEIVOLRSEVDHLRREITEREMQLTSSQKQTM 639

RESULT 10  
 US-09-864-761-47959  
 ; Sequence 47959, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
 ; FILE REFERENCE: Aecomica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47959
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AJ010770.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: Q99323, EVALUE 3.00e-17
; OTHER INFORMATION: EST_HUMAN HIT: AUI32932.1, EVALUE 1.00e-105
US-09-864-761-47959

Query Match          9.0%; Score 323; DB 10; Length 660;
Best Local Similarity 22.5%; Pred. No. 1.6e-07;
Matches 161; Conservative 135; Mismatches 238; Indels 180; Gaps 30;

QY 67 RYKVSSE---SKKESQKNDKILKEKTRVYLQERGAQDRRQDTELEKMEARLNA 122
DB 4 KVVNSNKIEQKQETIEELNTRIIIEEKRTL-----ELKDKLTITADKLGLGEQ 54

QY 123 ALAREKTSLSANNAFLKQLETRTNELLKSKFSENGKNLRL--SLEMLKLNKRET 180
DB 55 QIVQK-----NQEIKNNKLELT-----NSKQKRSSEIKQLMGTVBELQKNHSDS 102

QY 181 KMRGMAKQEGMEMKLVQTRSEESQKIAQLQEGK-----LVLSIEKEKIDE 227
DB 103 QFETDIVQR--ME---QETQKLEPQRAELDEMYGQIVQMKQELIRHQMAQMEKTRH 157

QY 228 KSETKILLEYIERTSCADQVEKYKLDIALEENLKEKNDLILSKQSLBNIVILSKQV 287
DB 158 KGMENALRSYNTIVNEDQIKLNVAINELNIKLOPTN----SQEKLKKEELGLI---- 209

QY 288 EDLNVKCOLLEKEDKHVNRNHNENLNAEMNLKQKFILEQOEHEKLOQKE---LQID 344
DB 210 --LEEK-ALRQLEDLVEELSFREQIARQVIAQESKLNKHAHKSLSVTDLKAIIV 266

QY 345 SLLQOEKELSSHLQKCSFQ---FEMVKEKNLFEELKQTLID-ELDKLOOK-----BEQ 395
DB 267 SASERKELEKLKHAETVNYKIKLEMEKEKNAYLDRMAESQEAELRLRTQLLFSHEEE 326

QY 396 AERLVKQLEBEAKSRAELK-----LLEELKLGKEAEL-----BKSSAAHTQATLL 441
DB 327 LSKLKDELEHRIINIEKLKDLNGLIHQKQIDGLQNMESQKIETMQFEKDLNLTQKNQLI 386

QY 442 LQ-EKYSMSVQSL-----EDVTAQFESYKALTASEIEDLKLENSLQOEKAAGKNAEDV 495
DB 387 LEISKLDQKQSLVNSSEMTLQ-----INELQKEIILROEEKEKGTLEQEV 435

QY 496 QHOILATESNOEYVRMLLDLQTKSALKETEIKETITSVFLQKITDQNLQKQOE----- 549
DB 436 Q-----ELQKLTLELKQMK-----KENDLQKQFALEANSIL 470

QY 550 EDFRKQLED-----EGR-----KAEKENTTAELTEEI-----NKWRLLY 584
DB 471 KDKBKTLEMLKIHPTVPSQERLFLDISKSDSVYWEKEIEILIEENEDLKQOCIOQLN 530

```



Qy 78 SQKNDKDKLEKEIRVLLQERGAQDRRIQDLETELEKMEARLNA--LRE 126  
Db 38 SRLSDK-IRLLQDDLEVERLQRIERKADLSVQVQIOMSERLFEABGGAHQPEANRKR 96  
Qy 127 KTSLSANNATLEKOLIELTRTNELLSKFSE--NGOKNRLILSLMLKLNKRKRETMRG 184  
Db 97 DAEULLKRLKLEDDHLESEETTLKLLKKHNEIITDFQEQVEILT-----KNK----- 143  
Qy 185 MMAQCEGMKQVOTORSLEESQKIAQLEGKLVISIEKIDKSETEKLEYLEETISCA 244  
Db 144 --ARAEDKAKFQT-----EYVELLSQTESYKNEKIVSKHISKLEVSISELVNK 191  
Qy 245 SDQVEKYKLDIAQLEENLKEKNDILSKOSLEENIIVLSKQVEDLVNKCOLLEKEKEDH 304  
Db 192 IEELNRTVIDISSHRSRLSQENIE-----LTKDVQDLKVLQDVTVSFSKSVQ 237  
Qy 305 VNRNREHNENLNAEMQNLKQKFILOQHEHEKLOOKELOIDSL----LOQRELSLSLHOKL 361  
Db 238 ISOLEDARRLEDED-----RRSLLESSLH-----QVEIELDSVRNQLSESEARIDLEROL 290  
Qy 362 CSFQEMVMEKKNLPEELKOTLDLDELKLOOK-----EEQABRLVKOLE--BEAKSR- 410  
Db 291 VKANADATSWQKNWSEVAPARAEEVEIRKRYQVRIETEHEHIESLIVKVNNEKMKTRL 350  
Qy 411 AEELKLL--BEKLGKREAELEKS-----SAAHTQATLLLOE-----KYDSM 449  
Db 351 ASEVEVLIDLEKSNNSCRELTQSVNTLKHNVKLKSLRDETIILYETSQORDLKNKHADL 410  
Qy 450 VOSLEDVTAQFESYKALTAS-----EIEDLKLENS-----S 480  
Db 411 VRTVHELDKVKDNNQNTRENKLGDDLHEAKGAINELNRLRLELELRLENERDELT 470  
Qy 481 LOEKAAGKAGNAEDVOHQILATESNQEVYVRMLLDLOTKSALKETEJ---KEITVSFLO 536  
Db 471 AAYKEAEGRAKAEQORQALADFNQYRH-----DAERLAEKDEEIEATRKQTSIEIQ 525  
Qy 537 ---KITDLQNLQOQEBDFRQLE-----DEEGRKAKEKENTTAETIEINKWRLLYEELYN 589  
Db 526 LNARVIEAETRLKTEVTRIKKLIQITELMSLDVANKTNIDLQKVIKQSLQTLTELQA 585  
Qy 590 KTKFPQLOLDAFEVKEKQALLNEHGAQEQOLNKIRDSYAKLLGHONLKQKIKHVYKLDEN 649  
Db 586 HYEDVQROL-----QATLDQYAVAQRRJLAGLNGELEEVRSHLDSANRAKRTVELQYE- 637  
Qy 650 SOLKSEYKLRQLAKKQSKQSTKLOEELNKV 680  
Db 638 -EASRINELTANVLSVSKLEQELSVV 667

## RESULT 14

US-09-815-242-13262  
; Sequence 13262, Application US/09815242  
; Patent No. US200200615691  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011a  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13262  
; LENGTH: 1179  
; TYPE: PRF  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13262

Query Match 8.7%; Score 314; DB 10; Length 1179;  
Best Local Similarity 21.9%; Pred. No. 6.4e-07;  
Matches 170; Conservative 141; Mismatches 265; Indels 202; Gaps 32;  
Qy 68 KVSSESSEKESOKNDKDKLEKEIRVLLQERGAQDRRIQDLETELEKMEARLNAALREK 127  
Db 171 KTRKETESKLDQQTODNLDRLIEDIYEL-----DNOIKLEKQAEANARFLDLEGQRK 223  
Qy 128 -----TSLSANNATLEKOLIELTRTNELLSKFSENGQKNLRLILSLMLKLNKRRE 179  
Db 224 VIYLDVLVAQIKENKAELESTEELAQVQELLMSVY-----QKREK-LEENQTLKKQORQ 277  
Qy 180 TKMRGMMAKQBGMEM-----KLOVTOBSL--EESQKIAQLEGKLVISIE 221  
Db 278 -DLQAEAKMQDQSGSLMDLTLSDLERKALSKLESEQVALNQEQEAQARLAALEDKRNLS 336  
Qy 222 KEKIDSETEKLEYLEETIEISCASQDQVEKYKLDIAQLEENLKEKNDILSLK----- 273  
Db 337 KEKIDKES-----SLALEGNLVQNNQKNLRLAEALLAFSD 372  
Qy 274 -----OSLEENIIVLSKQVEDLVNKCOLLEKEKEDH---VNRNREHNENLNAEMQNLK 324  
Db 373 DPQMIELLRRERFVALLQEEADVSNQLTRIENELNSQLSQKQADQLEKLEQOLATAKE 432  
Qy 325 KFILEQHEHEKLOQKELOIDSLLOQHEKLSLHOKLCSFOBEVMEKKNLEEBELKQTLQ 384  
Db 433 K--ASQCKDELETKVQVQKLLAD-----YQATAKEQEE--QKTSYQAQSQSLQD 478  
Qy 385 ELDKLOOKEQEAERLVKOLEEBAKSRABEELKLEK-----LKGKEAP-----L 428  
Db 479 RLDSLKNKQARAQSLNLRNHSNRYAGVKSVLQEKDKRGLGIIAGVSHLTFDYIYTAL 538  
Qy 429 EKSSAAHTQATLLQEKYDSMVQSLE-----DVTAFQESYKALT 467  
Db 539 EIALGASSQH--IIVEDEESATKAIDFLKRNVRGRATFLPTTIKARTISSQNQDAIAYS 596  
Qy 468 -----ASEIB--DLKLE-----NSSLOEKAAGKAGNAEDVOHQIL----- 500  
Db 597 PGFLGMADELVTFTRLAEAFKNNLATTAFDTVEHAREAAARQVRYQVRYVTLQGTLELT 656  
Qy 501 -----ATESNQEVYVRMLLD-LQTKSALKETEJ--KEITVSFLO---KITDLQNLK 546  
Db 657 GGSYAGGANRONNSIFIKPELEQLOKEITADEASLGSEEAALKTLQDQMAALTELEAIK 716  
Qy 547 QQEEEDFRKQLEDEGRKAKEKENTTAETIEINKWRLLYEEL-----YNTKTPFOLQ 597  
Db 717 SQEQARIQ---EQGLSLAYQQTQQVVELETMLKQEEIDRLSEGDQWQADKEK-CQES 772  
Qy 598 LDAFEVEKQALLNEHGAQEQOLNKIRDSYAKLLGHONLKQKIKHVYKLDKDE-NSQLKSEV 656  
Db 773 LATTASEKQNLAEETIEIKSNKNAIOERY-----QNLQEEVAQARLLKTLQLOGKRYEV 826  
Qy 657 SKLRCLQAKKQKQSETKLOEELNKVL-----GIKHFDPSKAFHESKENFALKTLPKEG 709  
Db 827 ADIE-RLGKELDNLNIQEEIQRMLQEKVDNLEKVD-TELLSQQAESKTKQTNLQQG 882

## RESULT 15

US-09-815-242-13608  
; Sequence 13608, Application US/09815242  
; Patent No. US20020061569A1

## GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13608

; LENGTH: 1179

; TYPE: PRN

; ORGANISM: Streptococcus pneumoniae

US-09-815-242-13608

Query Match 8.6%; Score 309; DB 10; Length 1179;

Best Local Similarity 21.5%; Pred. No. 1e-06;

Matches 166; Conservative 149; Mismatches 267; Indels 190; Gaps 29;

Qy 68 KVKSESKESKOKNDKLEKEIKRVLQERGAQDRRIQDLELEKEWEARLNAALREK 127

Db 171 KTRRKETESKLOQTQDNLDRLEDIYEL-----DNQIKPLEKQAEARKEFLDLEGORK 223

Qy 128 -----TSLSANWATLEKQILTRTNELLSKFSSENGKQKRLRILSLMLKRLNRKE 179

Db 224 ATYLDVLVAQIKDKNAELDSTEELAQVOELLTSYV-----QKREK-LEENQTLKKQRQ 277

Qy 180 TKMRGMMAKQEGMEM-----KLQVTRSL--EESQKIAQLEGKLVISIE 221

Db 278 -DLQAEAMAKDQGSMDLTSLSLDRKALKSEQAVALNQAEQAARLATLEDKRSSL 336

Qy 222 KEKIDKSEKLEKLEYIEBISCASQVBEKLDIAOLEENLEKKNDEILSLK----- 273

Db 337 QEKSKES-----SLALLEGNLVQNNQKLRLEAEALLAFSD 372

Qy 274 -----QSLERNIVLSKQVEDLVNFKCOLLEKEKEDHVNRRNHNENLNAEMQNLKQKFI 327

Db 373 DPDMIELLRERFVALLQEEADVSNQLTRINELN-----SRQLSQKQADQLEKLKEQLA 428

Qy 328 LEQOEHEKLOQKELQIDSLLOEKELSSSLKCLCSFOEEMVKEKNLFEELKQTLDEL 387

Db 429 TAK-----EKASQOELETAKEQVOKLLAD-YQATAKEOE---QKTSYQAQOSQLFRDL 481

Qy 388 KLOQKEQAERLVKOLEEAKSRABELKLEEK-----LKGKEAE-----LEKS 431

Db 482 NLKNQARAQSLNLRNHSNFYAGVKVQLQEKDRLGGLIIGAVSEHLTFDMHYQTAL 541

Qy 432 SAAHTQATLLQOE-----KYDSMVQSLEDVTAQFESYKA 465

Db 542 LGASSOHIIVEDENAATKAIDFLKRNAGRATFLPLTTIKARTISSQNQDAIAVSPGLG 601

Qy 466 LTASEIE--DLKLE-----NSSLOEKAAGKNAEDVOHQIL----- 500

Db 602 M-ADELVTFTDRLEAIFKNLLATTAFDTVEHAREAAQVRYQVRWVTLDTGTELTGSGY 660

Qy 501 ---ATESNQBYVRMLLD-LQTKSALKETEI--KEITVSFLQ----KITDQNLQKQEE 550

Db 661 AGGANRQNSIFIKPELQLOKEIAEEASLDSEVALKTLDDEMARLUTESLEAIKSGOE 720

Qy 551 DFRKOLEDEEGRKAEKENTTAELTEEINKWRLLYEELYNKI-----KPFQLQDAFE 602

Db 721 QARIQ---EOGLFLAYOQTSQOVEELETMLKQLEEEIDRLSEGDWQADKENCORLAAIA 777

Qy 603 VEQALLNEHGAQOEOLNKRDSYAKLLGHQNLKOKIKHVVKLKD-NSQLKSEVSKLRC 661

Db 778 SKONLEAEIEEIKSNKNAIQERY-----ONLQELAQARLLKTELOGQKRYEVADIE- 830

Qy 662 QLAKKKQSETKLQELNKNVLGIKHFDPSKA----FHESKENFALKTPLEK 709

Db 831 RLKELDNLNTEQEIEIQRMLOEKVDNLEKVDTLNQQAAEEAKTKTNLQOG 882

Search completed: December 20, 2002, 15:14:40

Job time : 16.0892 secs





```
QY 224 KIDEKSETEKLELEYIEISCASDOVEKYKLDIAQLEENLKEKNDILSKQSEENIVL 283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 KIDEKSETEKLELEYIQEISCASDOVEKCKVDIAQLEEDLKEKDRILSKQSEENIT-F 119
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 284 SKQVEDLVKCOLLEKEKEDHVNREHNENLNAEMQNLKQFIFLEQOHEKLOQKELOI 343
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 SKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQVEYKLOQKELO 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 344 DSLLOQKEKELSSSLHOKLCSFOEEMVKEKNLFEELKQTLDELKLOOKEQEAERLVKOL 403
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 QSLLOQKEKELSLARLOQLCSFOEEMTSEKNVFEELKALAEALDAVQOKEQSERLVKOL 239
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 404 EBEAKSRAEELKLEKLGKAELEKSSAAHTQATLL----- 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 EBEKSTAEQLTRLDNLREKEVELEKHTAAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 299
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 ----- 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 QEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQL 359
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 --LOEKYDSMWQSLQEDVTAQFESYKALTASETEDLKLENSLQEKAAKAGNAEDVOHQI 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 360 ESVQEKYNDTAQSLRDVTAQLESYKSSSTLKEIEDLKLENLTLOEKVMAEKSVEDVQOOI 419
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 500 LATESNQEVYRMLDLQTKSALKETEITEITVSFLQKITDQNLKQOEEDEFKQLEDE 559
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 LTAESTNQEVYARMVDLQNRSTLKEEKEITSSFLEKITDQNLKQOODEDFRQLEBK 479
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 560 EGRKAENKNTAELEENKRWLLYEELYNKTKPFQQLDAFEVEKQALLNEHGAQEO 619
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 GKRTAEKENVMTELTMEINKRWLLYDELYEKTQFQQLDAFEVEKQALLNEHGAQEO 539
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 620 NKIRDSYAKLLGHQNLKQIKHVVVKLDKNSQLKSEVSKRLCOLAKKQSETKLOEELNK 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 540 NKIRDSYAKLLGHQNLKQIKHVVVKLDKNSQLKSEVSKRLCOLAKKQSETKLOEELNK 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 680 VLGIKHFDPKAFHESKENFALKTPKLGKNTNC 713
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 600 ALGIRHFDPSKAFCHASKENF---TPLKEGPNPC 630
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
US-08-477-831C-2
; Sequence 2, Application US/08477831C
; Patent No. 6429291
; GENERAL INFORMATION:
; APPLICANT: TURLEY, EVA A.
; APPLICANT: SHUWEN, ZHANG
; APPLICANT: ENTWISTLE, JOYCELYN
; TITLE OF INVENTION: HYALURONAN RECEPTOR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10020-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Rel. #1.0, ASCII
; CURRENT APPLICATION DATA: US/08/477.831C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PIERRE, MARGARET A.
; REGISTRATION NUMBER: 30,709
; REFERENCE/DOCKET NUMBER: SIM-10
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: /desc = "RHAMM I protein"
US-08-477-831C-2

Query Match 50.6%; Score 1819.5; DB 4; Length 606;
Best Local Similarity 61.0%; Pred. No. 3.4e-95;
Matches 387; Conservative 50; Mismatches 84; Indels 113; Gaps 4;

QY 164 LRLSLELMKLRNKRRETKMRGMMAKQEGHEMKLVQTSRLESQGIKQIAQLEKLVSIK 223
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MRALSLELMKLRNKRRETKMRGMMAKQEGHEMKLVQTSRLESQGIKQIAQLEKLVSIK 54
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 224 KIDEKSETEKLELEYIEISCASDOVEKYKLDIAQLEENLKEKNDILSKQSEENIVL 283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55 -----CASDOVEKCKVDIAQLEEDLKEKDRILSKQSEENIT-F 94
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 284 SKQVEDLVKCOLLEKEKEDHVNREHNENLNAEMQNLKQFIFLEQOHEKLOQKELOI 343
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 SKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQVEYKLOQKELO 154
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 344 DSLLOQKEKELSSSLHOKLCSFOEEMVKEKNLFEELKQTLDELKLOOKEQEAERLVKOL 403
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155 QSLLOQKEKELSLARLOQLCSFOEEMTSEKNVFEELKALAEALDAVQOKEQSERLVKOL 214
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 404 EBEAKSRAEELKLEKLGKAELEKSSAAHTQATLL----- 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 EBEKSTAEQLTRLDNLREKEVELEKHTAAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 274
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 ----- 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 275 QEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQL 334
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 --LOEKYDSMWQSLQEDVTAQFESYKALTASETEDLKLENSLQEKAAKAGNAEDVOHQI 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 335 ESVQEKYNDTAQSLRDVTAQLESYKSSSTLKEIEDLKLENLTLOEKVMAEKSVEDVQOOI 394
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 500 LATESNQEVYRMLDLQTKSALKETEITEITVSFLQKITDQNLKQOEEDEFKQLEDE 559
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 LTAESTNQEVYARMVDLQNRSTLKEEKEITSSFLEKITDQNLKQOODEDFRQLEBK 454
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 560 EGRKAENKNTAELEENKRWLLYEELYNKTKPFQQLDAFEVEKQALLNEHGAQEO 619
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 455 GKRTAEKENVMTELTMEINKRWLLYDELYEKTQFQQLDAFEVEKQALLNEHGAQEO 514
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 620 NKIRDSYAKLLGHQNLKQIKHVVVKLDKNSQLKSEVSKRLCOLAKKQSETKLOEELNK 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 NKIRDSYAKLLGHQNLKQIKHVVVKLDKNSQLKSEVSKRLCOLAKKQSETKLOEELNK 679
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 680 VLGIKHFDPKAFHESKENFALKTPKLGKNTNC 713
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 575 ALGIRHFDPSKAFCHASKENF---TPLKEGPNPC 605
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-08-402-217A-2
; Sequence 2, Application US/08402217A
; Patent No. 5587301
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/402,217A  
FILING DATE: 10-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Luther, Barbara J.  
REGISTRATION NUMBER: 33954  
REFERENCE/DOCKET NUMBER: PF-00280US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-852-0195  
TELEFAX: 415-852-0195  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-402-217A-2

Query Match 47.1%; Score 1694; DB 1; Length 351;  
Best Local Similarity 98.9%; Pred. No. 2.2e-88;  
Matches 345; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 319 MQLKQKFILOQKEHEKLOQKELQIDSLLOQKEKELSSSLHOKLCSFQEMVKEKNLFEE 378  
Db 1 MQLKQKFILOQKEHEKLOQKELQIDSLLOQKEKELSSSLHOKLCSFQEMAKENLFEE 60  
Qy 379 LKQTLDELKLOQKEQERLVKQLEEEAKSRAEELKLEELKKEAELEKSSAAHTQA 438  
Db 61 LKQTLDELKLOQKEQERLVKQLEEEAKSRAEELKLEELKKEAELEKSSAAHTQA 120  
Qy 439 TLLQEKYDSMVQSLDVTQAFESYKALTASEIEDLKLENSSLQEKAAKAGNAEDVQHQ 498  
Db 121 TLLLEEKYDSMVQSLDVTQAFESYKALTASEIEDLKLENSSLQEKVAKAGNAEDVQHQ 180  
Qy 499 ILATSSNQEYVRMLDLQTKSALKETEIKETVTSFLQKITDLQNLKQOEEDFRKQLED 558  
Db 181 ILATSSNQEYVRMLDLQTKSALKETEIKETVTSFLQKITDLQNLKQOEEDFRKQLED 240  
Qy 559 EGRKAKEKNTAELTEINRWRLLYEELYNTKPFQQLDAFEVEKQALLNEHGAQEQ 618  
Db 241 EGRKAKEKNTAELTEINRWRLLYEELYNTKPFQQLDAFEVEKQALLNEHGAQEQ 300  
Qy 619 LNKIRDSYAKLLGHONLKQIKHVVKLKDENSQKSEVSKLRCLAKKK 667  
Db 301 LNKIRDSYAKLLGHONLKQIKHVVKLKDENSQKSEVSKLRCLAKKK 349

RESULT 4  
US-08-700-178-2  
Sequence 2, Application US/08700178  
Patent No. 5783669  
Patent No. 5783669 5700912  
GENERAL INFORMATION:  
APPLICANT: Hawkins, Phillip R.  
APPLICANT: Wilde, Craig G.  
APPLICANT: Seilhamer, Jeffrey J.  
TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN  
TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,178  
FILING DATE: August 20, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/402,217  
FILING DATE: March 10, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0028-1 DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-700-178-2  
Query Match 47.1%; Score 1694; DB 1; Length 351;  
Best Local Similarity 98.9%; Pred. No. 2.2e-88;  
Matches 345; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 319 MQLKQKFILOQKEHEKLOQKELQIDSLLOQKEKELSSSLHOKLCSFQEMVKEKNLFEE 378  
Db 1 MQLKQKFILOQKEHEKLOQKELQIDSLLOQKEKELSSSLHOKLCSFQEMAKENLFEE 60  
Qy 379 LKQTLDELKLOQKEQERLVKQLEEEAKSRAEELKLEELKKEAELEKSSAAHTQA 438  
Db 61 LKQTLDELKLOQKEQERLVKQLEEEAKSRAEELKLEELKKEAELEKSSAAHTQA 120  
Qy 439 TLLQEKYDSMVQSLDVTQAFESYKALTASEIEDLKLENSSLQEKAAKAGNAEDVQHQ 498  
Db 121 TLLLEEKYDSMVQSLDVTQAFESYKALTASEIEDLKLENSSLQEKVAKAGNAEDVQHQ 180  
Qy 499 ILATSSNQEYVRMLDLQTKSALKETEIKETVTSFLQKITDLQNLKQOEEDFRKQLED 558  
Db 181 ILATSSNQEYVRMLDLQTKSALKETEIKETVTSFLQKITDLQNLKQOEEDFRKQLED 240  
Qy 559 EGRKAKEKNTAELTEINRWRLLYEELYNTKPFQQLDAFEVEKQALLNEHGAQEQ 618  
Db 241 EGRKAKEKNTAELTEINRWRLLYEELYNTKPFQQLDAFEVEKQALLNEHGAQEQ 300  
Qy 619 LNKIRDSYAKLLGHONLKQIKHVVKLKDENSQKSEVSKLRCLAKKK 667  
Db 301 LNKIRDSYAKLLGHONLKQIKHVVKLKDENSQKSEVSKLRCLAKKK 349  
RESULT 5  
US-08-995-654-2  
Sequence 2, Application US/08995654  
Patent No. 6025138  
GENERAL INFORMATION:  
APPLICANT: Hawkins, Phillip R.  
APPLICANT: Wilde, Craig  
APPLICANT: Seilhamer, Jeffrey  
TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN  
TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:













TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2101 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-466-390-4

Query Match 11.0%; Score 395.5; DB 1; Length 2101;  
Best Local Similarity 22.2%; Pred. No. 1.4e-14;  
Matches 184; Conservative 163; Mismatches 303; Indels 179; Gaps 30;  
QY 5 KAPLKRENDPSGAP-----SPGAY---DVKTLEVLK-----GPVS--- 37  
DB 153 KAPV-----PSTCSTTFPEELSPSHQAKREIRFELEQKVASSSGNNFLSGSPASPMGD 207  
QY 38 -----FKSORFKQKESKONLVNVDKDTTLPASARKYKSSKESKOKNDKDLKILEKE 91  
DB 208 ILQTPQFQMRRLKKQLADERSN-----RDELELELAENRKLKLTKEKDAQIAMQOR 257  
QY 92 I--RVLLQERGA---QDRIQDLETELEKWEARLNALREKTSLSANNATLEKQIETL 145  
DB 258 IDRLALLNEKOAASPLEPKELEELDRKNESTLMTLHETLKQCQDLKTEKSDMRKINQLS 317  
QY 146 RTNELLKSKFSENGKQNLRLSLEMLKRLNKRKTKMRGMMAKQEGMEMKLOVT---ORS 202  
DB 318 EENGDLSPKLEFAS--HLQQLQDALNELTEHSEKATQEWLEKQAELEKLSAALQDKKC 375  
QY 203 LEES-----QKIAQLEGKLVIEKEKIDKSET-----EKLLEYIEISCASQVVEKY 251  
DB 376 LEERNEILQGLKSLQLEHLSQLDNPPQKEGVLGDLVQLQETLKQEAATLAANNITLQ-- 433  
QY 252 KLDIAQLEENKKNDEILSKQLEENIVILSKOVEDLVNKKCOLLEKEKEKEDHVNRRH 311  
DB 434 -ARVEMLETERGQOEAKLLAERGHFEKQQLSSLLITDQSSISNLSQAEELEQASQAH 492  
QY 312 NENLNAEMNLKQKFI-----LEQOEHEK-----LOOKELQIDSLLOQ 349  
DB 493 GARLTAQVASLTSETLTNATIQQDQBLAQKQAEKQALQATLQOQOQASOGLRHQ 552  
QY 350 EKELSSSLHQKLCSPQEWKKNLFEEELKQTLDELQKQKEQAEERLVKQLEEEAKS 409  
DB 553 VEQLSSSLKQKEQQLKEVAEQEATRODHAQQLATAAEERASLRERDAALKQLEALEKE 612  
QY 410 RAEELKLEELKQKEAELEKSSAAHTQATLLQEKYDSMVQSLDY-----TAQFESYK 464  
DB 613 KAAKLEILQQL---QVANEARDSAQTSVTQAREKAE--LSRKVEELQACVETARQEQHE 668  
QY 465 ALTASEIEDLKNSSLOEKAACKNAED---VHQHILATNESSNQEVYRMLDLQTKSA 521  
DB 669 A--QAQVALEQLRSEOKKATKERVAAQKQDQLOEQLOALKES--LKVTGKSLEEKERRA 725  
QY 522 LKTEIKETIVTSFKITDQNLQKQBEDFRKQLEDEP-----EGRKAEK 566  
DB 726 ADALAEQORCISLKAET---RSLVEQHKRKELEEEERAGKGLAEARLLQGEAQAET 782  
QY 567 ENTAEILTE-----EINKWRLLYEELYNKTKP-----FQQLDAFE--- 602  
DB 783 EVLRRELAEMAAQHTAESECEQLVKEVAARWDGYEDSQEEAQYGAMFQQLMTLKEC 842  
QY 603 -----VEKQALLNEHGAQ--EQLNKIROSYAKLL--GHQNLKQIKHVVYKLKDN 649  
DB 843 EKARQELQEAKEKQVAGTSHSELQISROQNKLAELHANLALQOQVQKEVRAQKLADDL 902  
QY 650 SOLK-----SEYSKRLCQAKKQSETKLQEEELNKLVLGKIKHFDPSKA 691  
DB 903 STLQEKMAATSEKVARLETLVKAGEQOQETASRELVK-----EPARA 944

RESULT 15

US-08-470-950-4  
; Sequence 4, Application US/08470950  
; Patent No. 5698439  
; GENERAL INFORMATION:  
; APPLICANT: TOUTKATLY, GARY  
; APPLICANT: LIDGARD, GRAHAM P  
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE  
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: 125 HIGH STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,950  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PITCHER ESQ, EDMUND R  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: MTP-013  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2101 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-470-950-4

Query Match 11.0%; Score 395.5; DB 1; Length 2101;  
Best Local Similarity 22.2%; Pred. No. 1.4e-14;  
Matches 184; Conservative 163; Mismatches 303; Indels 179; Gaps 30;

QY 5 KAPLKRENDPSGAP-----SPGAY---DVKTLEVLK-----GPVS--- 37  
DB 153 KAPV-----PSTCSTTFPEELSPSHQAKREIRFELEQKVASSSGNNFLSGSPASPMGD 207  
QY 38 -----FKSORFKQKESKONLVNVDKDTTLPASARKYKSSKESKOKNDKDLKILEKE 91  
DB 208 ILQTPQFQMRRLKKQLADERSN-----RDELELELAENRKLKLTKEKDAQIAMQOR 257  
QY 92 I--RVLLQERGA---QDRIQDLETELEKWEARLNALREKTSLSANNATLEKQIETL 145  
DB 258 IDRLALLNEKOAASPLEPKELEELDRKNESTLMTLHETLKQCQDLKTEKSDMRKINQLS 317  
QY 146 RTNELLKSKFSENGKQNLRLSLEMLKRLNKRKTKMRGMMAKQEGMEMKLOVT---ORS 202  
DB 318 EENGDLSPKLEFAS--HLQQLQDALNELTEHSEKATQEWLEKQAELEKLSAALQDKKC 375  
QY 203 LEES-----QKIAQLEGKLVIEKEKIDKSET-----EKLLEYIEISCASQVVEKY 251  
DB 376 LEERNEILQGLKSLQLEHLSQLDNPPQKEGVLGDLVQLQETLKQEAATLAANNITLQ-- 433  
QY 252 KLDIAQLEENKKNDEILSKQLEENIVILSKOVEDLVNKKCOLLEKEKEKEDHVNRRH 311  
DB 434 -ARVEMLETERGQOEAKLLAERGHFEKQQLSSLLITDQSSISNLSQAEELEQASQAH 492  
QY 312 NENLNAEMNLKQKFI-----LEQOEHEK-----LOOKELQIDSLLOQ 349  
DB 493 GARLTAQVASLTSETLTNATIQQDQBLAQKQAEKQALQATLQOQOQASOGLRHQ 552  
QY 350 EKELSSSLHQKLCSPQEWKKNLFEEELKQTLDELQKQKEQAEERLVKQLEEEAKS 409





```
QY 61 TLPASARKYKSSKESKOKNDKDLKILKEIRVLVLOERGAQDRRIQDLETELEKMEARL 120
DB 61 TLLASAKAKTLYSKKESKOKNDKDKVRKLEKEIRVLLQERGTQDKRIQDMSELETEAKL 120
QY 121 NAALREKTSLSANNATLEKQILTELITRNELLKSKFSFENGNOKNLRLTSLLEMLKLNKRET 180
DB 121 NAAVTEKPSLSANNASLEKRLTELITRANELLKSKFSFEDAHQNMRLTSLLEMLKLNKRET 180
QY 181 KWRGMAKOGEMKIQVORSLEESOGKIAOLEGKLVSTIEKEKIDSEKTEKLLIYIEE 240
DB 181 KWRGMAKOGEMKELQATOKDLTESGKIVOLEGKLVSTIEKEKIDSEKTEKLLIYIEE 240
QY 241 ISCASDQVEKYKLDITIAQLEENLKEKNDKDEILSLKQSLSEENIVILSKOVEDLNKVCQLEKE 300
DB 241 ISCASDQVEKYKLDITIAQLEENLKEKNDKDEILSLKQSLSEENIVILSKOVEDLNKVCQLEKE 300
QY 301 KEDHVRNREHNENLNAEMONLKQKILFEOQBEHEKLOQKELQIDSLQOKEKELSSSLHOK 360
DB 300 RDLVSKDRERAEISLSAEMQVLTERRALEROEYKILQQLNELQSLSLQOKEKELSAHLQOQ 359
QY 361 LCSFQEMVKEKNLFEELKQTLDELDKLQOKEQAEERLVKQLEEBKSRABELKLEEK 420
DB 360 LCSFQEMTSERNVFKELKLALDELDAVQKKQSEKLVKQLEEBTKSTAQELRLDNL 419
QY 421 LKGEAELEKSSAAHTQATILLQOQYDSMVQSLDVTAFESYKALTASEIEDLKLENS 480
DB 420 LREKETELEKRTAAHAQATVIAQEKYSDTAQLTRDVTQALYESYKSSTLKEIEDLKLENLT 479
QY 481 LQEKAAKAGKNAEDVQOILATNESSNOEYVRLMLDLQTSALKETEIKETVTSFLOKIND 540
DB 480 LQEKVAMAEKRVEDVQOQILTAESTNOEYAKVVQDLQNSSTLKEAEIKETTSYLEKIND 539
QY 541 LQNLQKQEDFRKQLEDEBEGRAEKENTTAELTEINKRWLLYEELYNKTPFQQLQIDA 600
DB 540 LQNLQKQEDFRKQLEEBGAKTEKETAVTETLMEINKRWLLYEELFPKTPFQQLQIDA 599
QY 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDENSQKSLR 660
DB 600 FEAQKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDENSQKSLR 659
QY 661 COLAKKQKQSETKLOEELNKVLGTHKHPDPSKAFHESKENFALTKPLKEGNTNC 713
DB 660 SOLAKRKQNELRLQGLDRLGIRHFDPSKAFCHESKENVTLTPLKEGNPNC 712

RESULT 2
Q8VDR2 ID Q8VDR2 PRELIMINARY; PRT; 836 AA.
AC Q8VDR2;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Similar to hyaluronan mediated motility receptor (RHAMM).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021427; AAH21427.1; -.
KW Receptor.
SQ SEQUENCE 836 AA; 96670 MW; B9BDDA22BECA652 CRC64;

Query Match 72.9%; Score 2623; DB 11; Length 836;
Best Local Similarity 65.1%; Pred. No. 5.5e-86;
Matches 546; Conservative 66; Mismatches 97; Indels 130; Gaps 3;

QY 1 MSFPKAPLARNDPSCAPSGAYDVKTLEVLKGPVSFOKSFQKQESKONLWVDKDT 60
DB 1 MSFPKAPLARNDPSCAPSGAYDVKTSEATKGPVSFOKSFQKQESKONLWVDKDT 60
```

```
QY 61 TLPASARKYKSSKESKOKNDKDLKILKEIRVLVLOERGAQDRRIQDLETELEKMEARL 120
DB 61 TLLASAKAKTLYSKKESKOKNDKDKVRKLEKEIRVLLQERGTQDKRIQDMSELETEAKL 120
QY 121 NAALREKTSLSANNATLEKQILTELITRNELLKSKFSFENGNOKNLRLTSLLEMLKLNKRET 180
DB 121 NAAVREKTSLSANNASLEKRLTELITRANELLKSKFSFEDGHQNMRLSLEMLKLNKRET 180
QY 181 KWRGMAKOGEMKIQVORSLEESOGKIAOLEGKLVSTIEKEKIDSEKTEKLLIYIEE 240
DB 181 KWRGMAKOGEMKELQATOKDLTESGKIVOLEGKLVSTIEKEKIDSEKTEKLLIYIOE 240
QY 241 ISCASDQVEKYKLDITIAQLEENLKEKNDKDEILSLKQSLSEENIVILSKOVEDLNKVCQLEKE 300
DB 241 ISCASDQVEKYKLDITIAQLEENLKEKNDKDEILSLKQSLSEENIT-FSKQIEDLTVCQLETE 299
QY 301 KEDHVRNREHNENLNAEMONLKQKILFEOQBEHEKLOQKELQIDSLQOKEKELSSSLHOK 360
DB 300 RDLVSKDRERAEISLSAEMQVLTERRALEROEYKILQQLNELQSLSLQOKEKELSAHLQOQ 359
QY 361 LCSFQEMVKEKNLFEELKQTLDELDKLQOKEQAEERLVKQLEEBKSRABELKLEEK 420
DB 360 LCSFQEMTSERNVFKELKLALDELDAVQKKQSEKLVKQLEEBTKSTAQELKRLDNL 419
QY 421 LKGEAELEKSSAAHTQATILLQOQYDSMVQSLDVTAFESYKALTASEIEDLKLENS 480
DB 420 LREKETELEKRTAAHAQATVIAQEKYSDTAQLTRDVTQALYESQEKYNDTAQSLRDVTAQ 479
QY 442 ----- 441
DB 480 LESVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQKYNDAQSLRDV 539
QY 442 -----LQEKYDSMVQSLDVTAFESYKALTASEIEDL 474
DB 540 TQALSEQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESYKSSTLKEIEDL 599
QY 475 KLENSLQEKAAKAGKNAEDVQOILATNESSNOEYVRLMLDLQTSALKETEIKETVTSF 534
DB 600 KLENLTQEKVAMAEKSVEDVQOQILTAESTNOEYARMVQDLQNRSTLKEEIKETTSF 659
QY 535 LOKITDLQNLQKQEDFRKQLEDEBEGRAEKENTTAELTEINKRWLLYEELYNKTPK 594
DB 660 LEKITDLKQNLQKQEDFRKQLEEKRTAEKENVMTLMEINKRWLLYEELYEKTKPF 719
QY 595 QLQDAFEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDENSQKLS 654
DB 720 QQLDFAFEAEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLKDENSQKLS 779
QY 655 EYSKURCOLAKKKQSETKLOEELNKVLGTHKHPDPSKAFHESKENFALTKPLKEGNTNC 713
DB 780 EYSKLRSQVLKQKQNELRLQGLDRLGIRHFDPSKAFCHASKENF---TPLKEGNPNC 835

RESULT 3
Q92026 ID Q92026 PRELIMINARY; PRT; 476 AA.
AC Q92026;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Hyaluronan receptor.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-92299690; PubMed-1376732;
RA Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,
RA Cripps V., Austen L., Nance D.M., Turley E.A.;
RT "Molecular cloning of a novel hyaluronan receptor that mediates tumor
cell motility.";
RL J. Cell Biol. 117:1343-1350(1992).
```









[illegible]



```
Db 789 OYKQOOLS-SGNSIDQQ-----LQSTIISELSEKQELNDSKLIBEKOLQOQOEFDOOL 843
Qy 103 DRRIQ-DLETEKMEARLNAALREKTSLSANNATLEKOLIELTRTNELLKSKFSENGNQ 161
Db 844 NEKNQKHQDQLELEKQLOQOQEDVQDQNETNQSIENQOONLTK-----ENLNE 896
Qy 162 KNLRLSLMLKRNKRETKMRGMMAKQEGMEMKQVQRTORSLEESQGGKIAQLEGLVSTE 221
Db 897 K-----EQELLQNLQNOQIEKIQDFQOEFQKQNSINIELVNEKNEKLIQLOQDYDQLK 951
Qy 222 KE--KIDEKSETEKILEYIEETISCASQVQEKYKLDIAQLEENLKEKNDEILSLKQSLSEN 279
Db 952 QOQRNSDEKEND-LIE-----KENOLKSIQNELNQLIEKNESDHKEQOQLKQOQSIEND 1003
Qy 280 IYVLSKQVEDLVNKKCOLLEK-----EKEDHVNRRNREHNENLNAEMONLKQKFILEQ 330
Db 1004 LIEKENQIQOL--QSOLENRQOQQNQSLSEKQDQQLNQLIEKNQ--FDQKEQOQLKQOQSIEND 1060
Qy 331 --QEHEKIQOKELOLDSLLQOQEKELSSLHQKLSFQOEMVKEKNLFEELKOTLDLQK 388
Db 1061 LFEKENQIQOQLQNEQOQQNQSLSEKQDQQLNQLIEK--NESQKEQOQLKQOQSIEND- 1117
Qy 389 LOQKEQAEARLVQLEEEAKSRAE-----ELKLEEKLGKKEAELEKSSAAHTQATLLL 442
Db 1118 LIEKENQIQOQLQNEQOQQNQSLSEKQDQQLNQLIEK--NESQKEQOQLKQOQSIEND 1177
Qy 443 QEK---YDSM-----VQSLVEDVTAQFESYKALTASEIEDLKLNSLQEK 484
Db 1178 QDKQIEFDQLQTFNQFNKNDKDSQFIQLODQKQOQSIQ-----QDLNOLKQOENQEKQ 1233
Qy 485 AAKAGKNAEDVQHQIATVSESSNQEVYRMLLDIQTSAKTEKEITVTFQKIDTDLQ-- 542
Db 1234 LSEKDEKLIQSIQFQNEQEKQSLSEKQDQQLNQLNQLNDEQNEKQVQSEKDEKLQSI 1293
Qy 543 ----NOLKQOEDFRKQLEDEGRKAENKTAETAEINKWRLLYEEINKTKPFQLOL 598
Db 1294 QODLNOLKQOENQEKQOL-----SEKDEKLQSQOQDLNQ---LNDDQIKKNEKLG--- 1340
Qy 599 DAFEVEKQALLNEHGAQOPOLNKRIDSYAKLIGHONLQKIKHVVYKLDKENSOLKSE--V 656
Db 1341 ---EKEEQLLQOQFNDQOQSQQLKQLEKSEKENQLOQLKQOENQOQLKQOENQOQOQSENI 1397
Qy 657 SKLRQLAKKQOSE-----TKLQBEL-----NKVLGT-----KHFD 687
Db 1398 QQLKQOQLKQOQOQOQENNEKEIERLQIEBQLKQOQOQOQSELSNEIKIQTTFQOQED 1457
Qy 688 PSKAFHESKENFALK 703
Db 1458 --QLSHNRKSKDQLHLQ 1471

RESULT 13
Q9BE41 PRELIMINARY; PRT; 1940 AA.
AC Q9BE41;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Myosin heavy chain 2a.
GN MYHC-2A.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN; TISSUE=SKELETAL MUSCLE;
RA Chikuni K., Muroya S., Nakajima I.;
RT "Sequencing of the bovine myosin heavy chain isoforms.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB059398; BAB40920.1; -;
```

```
DR HSSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; myosin_N; 1.
DR Pfam; PF01576; myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 1.
SQ SEQUENCE 1940 AA; 223318 MW; EE2642ELE29FDCC2 CRC64;

Query Match 11.5%; Score 414; DB 6; Length 1940;
Best Local Similarity 23.4%; Pred. NO. 2.le-07;
Matches 205; Conservative 135; Mismatches 261; Indels 276; Gaps 36;

Qy 51 KQNLNVKDXDTTLPASARKVKSESKEQKNDKDLKILEKEIRV-----LLQ 97
Db 724 KORYKVLNASAIPEG-----QYIDSKKSEKLLASIDIDHTQYKFGTKVFFKAGLLGLLE 779
Qy 98 ERGAQDRRIQDLETEKMEARLNAALREKTSLSANNATLEKQLELIELTRTNELLKSKFSE 157
Db 780 E--MRDEKLAQLMT---RTQARCRGFLAR-----VEYQKMWVERRESIEFC- 818
Qy 158 NGNQKNLR-----ILSLEMLKLNKRETKMRGMMAKQOQEMK--LQVQRTORSLEESQGKIA 211
Db 819 --IQYINRAFMNVKHPWNKLFRIKPLLSKSAETEKEMATKKEEPQKTDELAKSEAKRK 876
Qy 212 QLEGLVSTIEKIDEKSETKLELYIEISCASQOQEKYL-----DIAQLEENLKE 264
Db 877 ELEEKWVTLLEKNDLQLOVQSEAEGLADAEBCDQLIKTIQLEBAKKEVTERADEEE 936
Qy 265 KNDETLSLKQSLSEENIVILSKQVEDLVNKKQLEKEKEDHVNRRNREHNENLNAEMONL-- 322
Db 937 INAEUTAKKRLDEBCSELKDDIDDLLETLAKVEKEK--HATENKV--KNLTEEMAGLDE 992
Qy 323 -----KQKFILE-----QOHEKQLQ-----QKELQIDSL---LQOKELESS 356
Db 993 TIAKLTKEKKAQEAHQOQTLDDLAQAEEDKVNTLTAKTKLEQOQVDDLEGSLEQEKKLKRD 1052
Qy 357 LHO-----KLCSFOEMV---KEKNLPEELKQTLDELQKQK--EBOA----- 396
Db 1053 LERAKRKLEGLDKLA--QESIMDINENKQQLDEKLKKEFEISNLSQSKIEDQALGIQLQ 1110
Qy 397 -----ERLVKQLEEEAKSRAE-----ELKLLPEKL-----K 422
Db 1111 KKIKELQARIEELEEIEAERASRAKAEKQSDLSRELEEEISERLEEAGGATSAQIEMNK 1170
Qy 423 KEAEB-----LEKSSAAHTQATLLQEKY--DSM-----VQSLVEDVTAQFESYKALTA 468
Db 1171 KREAEFQKMRDLEPATLQHEATAALRRKHADSVAEELGEQIDNLQRYKQLEKEKSEMK 1230
Qy 469 SEIEDKLKLENSLOEKAACKNAEDVOHQIATVSESSNQEVYRMLLDIQTSAKTEI- 527
Db 1231 MEIDDLASNVETISKAKGNLEKMCRTLEDQVNELSKSEEEQORLINDLTQORGLQTESG 1290
Qy 528 -----KEITVS-----FLQKITDQLQNLQKQ-----EED---PRK 554
Db 1291 EFSRQDKEALVSQLSRGKQAFQOIEELKQLEEEIKAKNALAHQLSARHDCDLLRE 1350
Qy 555 QLEDEEGRKAKENTTAELTEINKWRLLY-----EELYNKTKPQLQIDA----- 600
Db 1351 QYEEQESKAEQRLSKANTEVAQWRKTYETDAITQRTTEELEEAQKKAQRLQAEEHVE 1410
Qy 601 -----FEVEKQALLNE-----HGAQOQLNKRIDSYAKLIG-----HQ 633
Db 1411 AVNAKASLEKTKORLQNEVEDLMDVETNAACALDKORNFDKILAEWKQKYETHA 1470
Qy 634 NL-----KQIKHVVYKLDKENSQSKSEVSKRLCQLAK----- 665
```

```
Db 1471 ELEAAKREARSLGTFLFKMKNAYEESLDQLETLKRENKNLQOEISDLTEQIAEGCKRMHE 1530
Qy 666 ----KKQ-----SETKQELNKNVGLGK 684
Db 1531 LEKIKKQVEQKSEIQAALEAEASLEHEGKILRIQ 1567

RESULT 14
Q922D2
ID Q922D2 PRELIMINARY; PRT; 1598 AA.
AC Q922D2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to myosin, heavy polypeptide 2, skeletal muscle,
DE adult.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008538; AAH08538.1; -.
DR InterPro; IPR000048; IQ.region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR PROSITE; PS00096; IQ; 1.
SQ SEQUENCE 1598 AA; 183083 MW; DE896B34346A479 CRC64;

Query Match 11.5%; Score 413; DB 11; Length 1598;
Best Local Similarity 24.8%; Pred. No. 1.9e-07;
Matches 184; Conservative 114; Mismatches 205; Indels 240; Gaps 30;

Qy 172 MKLRNRETKRGMMAKQEMEMK--LQVTRSLSESOCKTAQLRGKLVSEKEKIDKS 229
Db 837 MKLFPKIRPLKSAETEKEMATMKKEFQTKDDLAKSEAKRKELEKVVSLLEKNDLQL 896
Qy 230 ETEKLEVEIEISASDOVEYKL-----DIAOLENLEKNDLILSLKQSLLENIVI 282
Db 897 QVQAEAGLADAEERCDQLIKTKIOLEAKIKEVTERAEDEEIEINAEALTAKRKLEDCSE 956
Qy 283 LSKOVEDLNVRCOLLEKEKEDHVNREHNENLNAMQNL-----KQKFILE----- 329
Db 957 LKKDIDDLLETLAKVEKEK--HATENKV--KNLTTEAGLDETIAKLTKERKAQEAHQ 1012
Qy 330 -----QOEHEKIQ-----QKEQLIDSL--LQOEKELSSLSHQ-----KLCFQ 365
Db 1013 TLDDLQAEEDKVNLTAKIKLEQQVDDLEGSLEQKLLWDLERAKRKLEGLDLA--Q 1070
Qy 366 EMV---KEKNLFEELKOTDELKLOOK--EEQA-----ERLVKOLE 404
Db 1071 ESIMDIENEKQOLDERLAKKEFEMSNLQSKIDEQAIGTQKQKIKELQARTEEEETE 1130
Qy 405 EAKSRAE-----ELKILLEKL-----KGKAE-----LEKSSA 433
Db 1131 AERASRAEAKORSOLSLRELEEISERLEEAGATSAQIEMNKKRAEAFQKMRDLLEAATL 1190
Qy 434 AHTQATLLQEKY--DSM-----VQSLDVTQAQFYSKALTASEIEDLKLENSLQEKAA 486
Db 1191 QHEATAATLRKHDVSLAEGLGEQIDNLRVKQKLEKSEKMEIDDL-----ASNVEIVS 1246
Qy 487 KAGKNAE----DVQHQIILATESNQYVRMLDLDTKSAKRETEI-----KEITVS 533
Db 1247 KAKGNLEKMCRTLDQVSELKSEEEQOQLINDLTSQGRQLQTESGEFSRQDKEALVS 1306
Qy 534 -----FLOKITDLOLKOQOED-----FRKQLEDESGRAEKEN 568
Db 1307 QLSRGKQAFQOIEELKQLEEEVAKNALAHALOSSRHDCCDLLREQYEEESKAELQR 1366
Qy 569 TTAELTEEINKWRLLY-----BELYNKTKPFOQLDA-----FEVEKQ 606
Db 1367 ALSKANSEVAQWRKYETDAIQORTEELBEAKKLAQRLQAEEHVEAVNAKCASLEKTKQ 1426
Qy 607 ALLNE-----HGAQAQELNKRDSYAKLLG-----HQNLL----- 635
Db 1427 RLQNEVEDLMLDVERTNAACAALDKQRNDFKILAEWKQYETHAELEASKEARSIGT 1486
Qy 636 -----KQIKHVVKLKDENSOLSKSEVSKLRQLAK-----KKQ----- 668
Db 1487 ELFKMKNAYEESLDQLETLKRENKNLQOEISDLTEQIAEGGKRIHELEKIKKQVEKCE 1546
Qy 669 -----SETKQELNKNVGLGK 684
Db 1547 LQAALAEAEASLEHEGKILRIQ 1569

RESULT 15
Q9NS87
ID Q9NS87 PRELIMINARY; PRT; 1388 AA.
AC Q9NS87;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Kinesin-like protein 2.
GN HKLP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20435852; PubMed=10878014;
RA Sueishi M., Takagi M., Yoneda Y.;
RT "The forkhead-associated domain of Ki-67 Antigen Interacts with the
RT Novel Kinesin-like Protein Hk1p2.";
RL J. Biol. Chem. 275:28888-28892(2000).
DR EMBL; AB035898; BAB03309.1; -.
DR HSP; p17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 1388 AA; 160160 MW; E127EB4B9991CA83A CRC64;

Query Match 11.5%; Score 412; DB 4; Length 1388;
Best Local Similarity 21.0%; Pred. No. 1.8e-07;
Matches 171; Conservative 157; Mismatches 282; Indels 206; Gaps 25;

Qy 37 SFQKQSQREKQ-----QKESKNLVND-----K 58
Db 598 SKQEEFEKELTRKQLELESELSQKLANLENLEATKACKRQEVSQLNKHIAETLK 657
Qy 59 DTTLPASARKVKSSSEKES-----QKNDKDLKLEK 90
Db 658 ITTTTKAYOLHSRVPVKLSPWGSFGSLYTONSSITLDNILENPVPPENQEAPEATSE 717
Qy 91 EIRVLQERGAODRIQD-----LETELEKME-----ARUNAAALREKTSUS 131
Db 718 ELURTVOEQMSALQALDDEEHKNLQHQVHDKLEHHSTOMQSELSFSSERIDMTKQOEELIS 777
Qy 132 ANNATLEQOLIELTNTNELLKSKFSE-----NGNQKNLILSLMLKLNKRETKRMGM 186
Db 778 QLN-VLEKQLQETQTKNDNDFLKVSEVHDLRVVLHSADEKELSSVKLEYSSSFKTNOKEFENKL- 835
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:06:53 ; Search time 10.6932 Seconds  
(without alignments)  
2812.095 Million cell updates/sec

Title: US-09-685-010-47

Perfect score: 3598

Sequence: 1 MSFPKAPLKRFPNDPGCAPS.....LKEGNTNCRAPMEQCQESWK 725

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3568.5	99.2	724	1	HMMR_HUMAN
2	2623	72.9	794	1	HMMR_MOUSE
3	1768.5	49.2	498	1	HMMR_RAT
4	444	12.3	1790	1	USO1_YEAST
5	422	11.7	2230	1	GOG4_HUMAN
6	419	11.6	1976	1	MYHA_RAT
7	417	11.6	1875	1	MLP1_YEAST
8	414	11.5	1976	1	MYHA_HUMAN
9	414	11.5	2116	1	MYSD_DICDI
10	409.5	11.4	1939	1	MYH4_HUMAN
11	409	11.4	1959	1	MYH9_CHICK
12	408	11.3	2017	1	MYSN_DROME
13	407.5	11.3	1433	1	REST_CHICK
14	407	11.3	1941	1	MYH2_HUMAN
15	405	11.2	1679	1	YIO9_YEAST
16	404.5	11.2	1960	1	MYH9_HUMAN
17	403.5	11.2	1938	1	MYSS_CHICK
18	403	11.2	1938	1	MYH4_RABIT
19	403	11.2	1940	1	MYH3_RAT
20	402.5	11.2	1972	1	MYHB_RABIT
21	402	11.2	1978	1	MYHB_CHICK
22	400.5	11.1	1940	1	MYH3_CHICK
23	398.5	11.1	1972	1	MYHB_HUMAN
24	398	11.1	1102	1	MYSC_CHICK
25	398	11.1	1940	1	MYH3_HUMAN
26	398	11.1	1972	1	MYHB_MOUSE
27	398	11.1	1976	1	MYHA_BOVIN
28	396	11.0	1937	1	MYH8_HUMAN
29	395.5	11.0	1427	1	REST_HUMAN
30	395	11.0	1961	1	MYH9_RAT
31	394.5	11.0	1938	1	MYHD_HUMAN
32	393	10.9	976	1	SCP1_HUMAN
33	393	10.9	1939	1	MYH1_HUMAN

#### RESULT 1

ID	HMMR_HUMAN	STANDARD;	PRT;	724 AA.
AC	075330; Q92767;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Hyaluronan mediated motility receptor (intracellular hyaluronic acid binding protein) (Receptor for hyaluronan-mediated motility) (CD168 antigen)			
DE	HMMR OR IHABP OR RHAMM.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION. TISSUE=Breast carcinoma;			
RX	MEDLINE=98264864; PubMed=9601098;			
RA	Assmann V., Marshall J.F., Fieber C., Hofmann M., Hart I.R.;			
RT	"The human hyaluronan receptor RHAMM is expressed as an intracellular protein in breast cancer cells.";			
RL	J. Cell Sci. 111:1685-1694(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Breast;			
RX	MEDLINE=97045829; PubMed=8890751;			
RA	Wang C., Entwistle J., Hou G., Li Q., Turley E.A.;			
RT	"The characterization of a human RHAMM cDNA: conservation of the hyaluronan-binding domains.";			
RL	Gene 174:299-305(1996).			
CC	-!- FUNCTION: INVOLVED IN CELL MOTILITY. WHEN HYALURONAN BINDS TO HMMR, THE PHOSPHORYLATION OF A NUMBER OF PROTEINS, INCLUDING THE FOCAL ADHESION KINASE OCCURS. MAY ALSO BE INVOLVED IN CELLULAR TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY.			
CC	-!- SUBUNIT: SUBUNIT OF THE HARC COMPLEX.			
CC	-!- SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC (BY SIMILARITY).			
CC	-!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN BREAST CANCER CELL LINES AND IN NORMAL BREAST TISSUE.			
CC	-!- DATABASE: NAME=PROW; NOTE=PROW 2:76-84(2001); WWW="http://www.ncbi.nlm.nih.gov/prow/guide/802868666.g.htm".			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; AF032862; AAC32548.1; -.			
DR	EMBL; U29343; AAC52049.1; -.			

P54697 dictyosteli  
Q62209 mus musculu  
P12883 homo sapien  
P13533 homo sapien  
P13539 mesocricetu  
Q90339 cyprinusc ca  
Q02566 mus musculu  
Q03410 rattus norv  
O67124 aquifex aeo  
Q10411 schizosacch  
P49454 homo sapien  
Q99996 h a-kinase

DR Genew: HGNC:5012; HMMR.

DR MIM: 600936; -.  
 KW Hyaluronic acid; Alternative splicing; Repeat; Glycoprotein; Antigen.  
 FT DOMAIN 635 645 HYALURONIC ACID-BINDING (POTENTIAL).  
 FT DOMAIN 657 666 HYALURONIC ACID-BINDING (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 588 588 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 75 90 MISSING (IN ISOFORM B).  
 FT CONFLICT 75 75 K -> KK (IN REF. 2).  
 FT CONFLICT 103 103 S -> R (IN REF. 2).  
 FT CONFLICT 277 277 E -> D (IN REF. 2).  
 FT CONFLICT 298 298 K -> T (IN REF. 2).  
 FT CONFLICT 322 322 K -> E (IN REF. 2).  
 FT CONFLICT 330 332 OER -> REH (IN REF. 2).  
 SQ SEQUENCE 724 AA; 84031 MW; EA68AD6D2A626926 CRC64;

Query Match 99.2%; Score 3568.5; DB 1; Length 724;  
 Best Local Similarity 99.6%; Pred. NO. 1.8e-109;  
 Matches 722; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MSFPKAPLRFNDPSCAPSGAYDVKTLEVLKGPVSPQKQSFQKQESKONLVNDKDT 60  
 Db 1 MSFPKAPLRFNDPSCAPSGAYDVKTLEVLKGPVSPQKQSFQKQESKONLVNDKDT 60  
 QY 61 TLPASARKVKSSEKQNDKDLKLEKEIRVLVLLQERGAQDRRIQDLETELEKWEARL 120  
 Db 61 TLPASARKVKSSES-KESQNDKDLKLEKEIRVLVLLQERGAQDSRIQDLETELEKWEARL 119  
 QY 121 NAALREKTSLSANNATLEKQLELTRTNELLKSKFSENGQKNRLSLELMKLNKRKET 180  
 Db 120 NAALREKTSLSANNATLEKQLELTRTNELLKSKFSENGQKNRLSLELMKLNKRKET 179  
 QY 181 KVRGMAKQEGMEMKQVTFQSRLESQKIAQLEKGLYSIEKEIDKSETEKILLEYTEE 240  
 Db 180 KVRGMAKQEGMEMKQVTFQSRLESQKIAQLEKGLYSIEKEIDKSETEKILLEYTEE 239  
 QY 241 ISCASDOVEKYKLDIAQLEENKEKNDLTSLSKQSEENIVILSKQVEDLNKVCOLLEKE 300  
 Db 240 ISCASDOVEKYKLDIAQLEENKEKNDLTSLSKQSEENIVILSKQVEDLNKVCOLLEKE 299  
 QY 301 KEDVNRNREHNENLNAEQNLKQFPILEQQBEKLOQKEIQIDSLLOQKEKELSSSLHOK 360  
 Db 300 KEDVNRNREHNENLNAEQNLKQFPILEQQBEKLOQKEIQIDSLLOQKEKELSSSLHOK 359  
 QY 361 LCSFOEMVKENLFEELKQTLDELKLOQKEQAEALYKOLEEAKSRAPEELKLEEK 420  
 Db 360 LCSFOEMVKENLFEELKQTLDELKLOQKEQAEALYKOLEEAKSRAPEELKLEEK 419  
 QY 421 LKGEAELEKSSAAHTQATLLLOEKYDSMVQSLQEDVTAQFESYKALTASETEDLKLENS 480  
 Db 420 LAGKAELEKSSAAHTQATLLLOEKYDSMVQSLQEDVTAQFESYKALTASETEDLKLENS 479  
 QY 481 LOEKAAGKAGNAEDVOHQILATPSSNQEVYRMLDLQTKSALKETEIKETVSPFLOKQITD 540  
 Db 480 LOEKAAGKAGNAEDVOHQILATPSSNQEVYRMLDLQTKSALKETEIKETVSPFLOKQITD 539  
 QY 541 LQNLKQOEEDPRKOLEDEEGKAKENTTAELTEINKRWLLYELYNKTKPFQLOLDA 600  
 Db 540 LQNLKQOEEDPRKOLEDEEGKAKENTTAELTEINKRWLLYELYNKTKPFQLOLDA 599  
 QY 601 FEVEKQALINEHGAQAQEQNLIRDSYAKLLGHONLKQIKHVVKLDNSOLKSEVSKLR 660  
 Db 600 FEVEKQALLNEHGAQAQEQNLIRDSYAKLLGHONLKQIKHVVKLDNSOLKSEVSKLR 659  
 QY 661 COLAKKKQSETKLQELNKVLGIKHFDPSPKAFPHHESKENFALKTPKKGNTNCRYAPMEC 720  
 Db 660 COLAKKKQSETKLQELNKVLGIKHFDPSPKAFPHHESKENFALKTPKKGNTNCRYAPMEC 719  
 QY 721 QESWK 725  
 Db 720 QESWK 724

RESULT 2  
 HMMR\_MOUSE STANDARD; PRT; 794 AA.  
 ID Q00547;  
 AC 01-OCT-1996 (Rel. 34, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Hyaluronan mediated motility receptor (intracellular hyaluronan acid binding protein) (Receptor for hyaluronan-mediated motility).  
 DE HMMR OR IHABP OR RHAMM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.  
 RC TISSUE=Lung;  
 RX MEDLINE=98264863; PubMed=9601097;  
 RA Hofmann M., Fieber C., Asmann V., Goettlicher M., Sleeman J., Plug R., Howells N., von Stein O., Ponta H., Herrlich P.;  
 RT "Identification of IHABP, a 95 kDa intracellular hyaluronate binding protein.";  
 RL J. Cell Sci. 111:1673-1684(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Zhao Y., Zhang S., Turley E.;  
 RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-183 FROM N.A., AND TISSUE SPECIFICITY.  
 RC STRAIN=129/SV;  
 RX MEDLINE=99107769; PubMed=9889313;  
 RA Fieber C., Plug R., Sleeman J., Dall P., Ponta H., Hofmann M.;  
 RT "Characterization of the murine gene encoding the intracellular hyaluronan receptor IHABP.";  
 RL Gene 226:41-50(1999).  
 RN [4]  
 RP SEQUENCE OF 164-794 FROM N.A., AND ALTERNATIVE SPLICING.  
 RC STRAIN=BALB/c; TISSUE=Fibroblast;  
 RX MEDLINE=96011639; PubMed=7590272;  
 RA Entwistle J., Zhang S., Yang B., Wong C., Li Q., Hall C.L., A J., Mowat M., Greenberg A.H., Turley E.A.;  
 RT "Characterization of the murine gene encoding the hyaluronan receptor RHAMM.";  
 RL Gene 163:233-238(1995).  
 RN [5]  
 RP SEQUENCE OF 318-794 FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=BALB/c;  
 RX MEDLINE=92299690; PubMed=1376732;  
 RA Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D., Cripps V., Austen L., Nance D.M., Turley E.A.;  
 RT "Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motility.";  
 RL J. Cell Biol. 117:1343-1350(1992).  
 RN [6]  
 RP CHARACTERIZATION.  
 RX MEDLINE=94308286; PubMed=7518470;  
 RA Hall C.L., Wang C., Lange L.A., Turley E.A.;  
 RT "Hyaluronan and the hyaluronan receptor RHAMM promote focal adhesion turnover and transient tyrosine kinase activity.";  
 RL J. Cell Biol. 126:575-588(1994).  
 RN [7]  
 RP ERK REGULATION, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=98252522; PubMed=9556628;  
 RA Zhang S., Chang M.C., Zylka D., Turley S., Harrison R., Turley E.A.;  
 RT "The hyaluronan receptor RHAMM regulates extracellular-regulated kinase.";  
 RL J. Biol. Chem. 273:11342-11348(1998).  
 RN [8]  
 RP REVIEW.  
 RX MEDLINE=99059494; PubMed=9845361;  
 RA Hofmann M., Asmann V., Fieber C., Sleeman J.P., Moll J., Ponta H.,







```
Db 1006 MSQKENFQIERGSIENKIEIQLKTTISLEQTKKEIISKSDSSKDEYESQISLLKEKLET 1065
Qy 129 SLSANNATLEKQELTERTNELLSKSENGNOKNRLISLEMLKRNKRETKMRGMMAK 188
Db 1066 ATTANDENVK-ISELTKTRELEAEALAAAYKLNK------ELETKLETSEKALKEV 1115
Qy 189 QEG-----MEMKQVQTSLEESQKIAQLEGLKYSIEKEKID-----EKSETEKILLE 236
Db 1116 KENEHLKEEKIQLEKEKATETKQ-QLNSLRANLESLEKEHEDLAQALKYEEQIANKERQ 1174
Qy 237 YIEISCASDOVEYKLDIAQLENLEKND-----EILSLKQSLNEENIVILSKQVEDLNV 292
Db 1175 YNEISQINDBITSTQ-----QENESIKKKNDLEGELEGKAMKSTSEEQSNLKKSDIDLNL 1230
Qy 293 KCOLLEKEKEDHVRNRHNNENNAEQNLKOKFILOQKEHEKLOQKLEQIDSLLOQKE 352
Db 1231 QIKELKKNETNEASLESIKSVESEVVK-----ELQDCNFEKEV-----SELEDKLK 1282
Qy 353 LSSSLHOKLCSFQEMVKEKNLFE---EELKQTLDELKLOQKEQAE---RLVKOLEE 405
Db 1283 ASEDKNSKYLEQKESEKIKBELDAKTTELKIQLEKITNLKSAREKSESELSRLKKTSE 1342
Qy 406 EAKSRAEKLLEKLKKEAELEKSSNAHTQATLLLOQKYDSMVQSLVEDVTAQESYKA 465
Db 1343 ERKNAEQLEKLKWEIQKNOAFKERRKLLNEGSTITQOEYSEKINTLEDELIRLQNE 1402
Qy 466 L-----TASEIDKLLENSSLOEAKAGKAGNAED----- 494
Db 1403 LKAKEIDNTRELEKVSLSNDELLEEQNTIKSLQDEILSYKDKITRNDKLLSIEDNKK 1462
Qy 495 -----VQHILATSSNOEYVRLMLDLOTKSALKETEI---KEITVSFLQKITDLQNLK 546
Db 1463 RDLSELAQLRAAOSKAKVEGLKLEESSEKAELEKSEMMKLESTIESNETLK 1522
Qy 547 QOEDFQKQLED-BEGRKAKE-----NTTAEITELINKWRLLYEELYNYTKFPQOLD 599
Db 1523 SSMETIRKSDKLSQSKSAEEDTKNQHEKSDLSIRISEKEDIELKSK-----LRIE 1577
Qy 600 A-----FEVEKQALNEHGAEOQLNKIRDSYAKLLHQH--LKQKIKHVYK-LKDENSQ 651
Db 1578 AKSGSELETVKQLNN---AQEKI-----RINAENIVLAKSLEDIERELKDKQAE 1625
Qy 652 LK-----SEVSKLRQALAKKQSETKLQBELKNVLGIKHPDFSPKAPHHESKENFA 701
Db 1626 IKSNOEKEKELLTSRLKELEQELDSTQQAOKSEEE--RRAEVKRFQVEKSLDE----- 1677
Qy 702 LKPLKCGTNCYRAPMEQCSWK 725
Db 1678 -KAMLETK-----YNDLVNKEQAWK 1697

RESULT 5
GOG4_HUMAN STANDARD; PRT; 2230 AA.
AC Q13439; Q14436; Q13270; Q13654;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Golgi autoantigen, golgin subfamily A 4 (Trans-Golgi p230) (256 kDa
DE golgin) (Golgin-245) (72.1 protein).
GN GOLGA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96215236; PubMed=8626529;
RA Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.;
RT "Molecular characterization of trans-Golgi p230: a human peripheral
RT membrane protein encoded by a gene on chromosome 6p12-22 contains
RT extensive coiled-coil alpha-helical domains and a granin motif.";
RL J. Biol. Chem. 271:8328-8337(1996).
```

```

[2]
RN SEQUENCE FROM N.A.
RP Seelig H.P.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE OF 131-2230 FROM N.A.
RP TISSUP-Placenta;
RX MEDLINE=96125112; PubMed=8537393;
RA Fritzler M.J., Lung C.-C., Hamel J.C., Griffith K.J., Chan E.K.L.;
RT "Molecular characterization of golgin-245, a novel Golgi complex
RT protein containing a granin signature.";
RL J. Biol. Chem. 270:31262-31268(1995).
[4]
RN SEQUENCE OF 524-672 FROM N.A.
RP TISSUP-Gastric fundus;
RA Balague C.;
RL Thesis (1994), Instituto municipal de Investigacion medica, Spain.
CC -!- FUNCTION: MAY PLAY A ROLE IN VESICULAR TRANSPORT FROM THE TRANS-
CC GOLGI.
CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC FACE OF THE
CC GOLGI MEMBRANE.
CC -!- ALTERNATIVE PRODUCTS: At least 4 isoforms; 1 (shown here), 2, 3
CC and 4; are produced by alternative splicing.
CC -!- DISEASE: ANTIGEN IN THE AUTOIMMUNE DISEASE SJOEGREN'S SYNDROME AND
CC IN HEPATITIS B.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U41740; AAC50434.1; -
CC EMBL; X82834; CAAS8041.1; -
CC EMBL; U31906; AAC51791.1; -
CC EMBL; X76942; CAAS4261.1; -
CC Genew; HGNC:4427; GOLGA4.
CC MIM; 602509; -
CC InterPro; IPR000237; GRIP_domain.
CC Pfam; PF01465; GRIP; 1.
KW Golgi stack; Antigen; Coiled
FT DOMAIN 133 237
FT COILED COIL (POTENTIAL).
FT DOMAIN 276 1011
FT COILED COIL (POTENTIAL).
FT DOMAIN 1033 1214
FT COILED COIL (POTENTIAL).
FT DOMAIN 1259 2152
FT COILED COIL (POTENTIAL).
FT VARSPLIC 2154 2185
FT TPYKGNLYHVDVSLFGEPTFEYLKRVLFY -> HLTKV
FT AICTIRMSHLENPLNLSICEKFLSI (IN ISOFORM
FT 2).
FT VARSPLIC 2186 2230
FT MISSING (IN ISOFORM 2).
FT VARSPLIC 2103 2109
FT MISSING (IN ISOFORM 3).
FT VARSPLIC 2222 2230
FT FTSPRGIF -> SWLRSS (IN ISOFORM 4).
FT R -> K (IN REF. 3).
FT Y -> H (IN REF. 3).
FT Y -> A (IN REF. 3).
FT CONFLICT 276 276
FT CONFLICT 584 584
FT CONFLICT 628 628
FT CONFLICT 630 630
FT CONFLICT 682 682
FT K -> N (IN REF. 3).
SQ SEQUENCE 2230 AA; 261139 MW; 3BB733DB1EA86134 CRC64;

Query Match 11.7%; Score 422; DB 1; Length 2230;
Best Local Similarity 22.8%; Pred. No. 3.1e-07;
Matches 216; Conservative 151; Mismatches 275; Indels 306; Gaps 39;

Qy 26 VKTLEVLKGPVSFQSKRFQKQ-----KESQNLNVYDKD-TTLPASARKVKSESSESQK 80
Db 278 VKTLETLQ-----QVRKQENLLKRRCKETIQSHKQCTILLTSEKALQQLDERLQ 329
Qy 81 ND--KDLKILKE-----IRVLQERG---AQDRR-----TDLTELEKWEA 118
Db 330 LEKIKDLHMAEKTLLITQLRDAKNLIEQLEQDKGMVIAETKQRMHETLEMKEEIAQLRS 389
```



```

QY 193 EMKLVQTSRLESOGKTAQLEGLKLVSTIEKEIDEKSETKLEYLEIEISCASQDQVEKYK 252
Db 1314 ESQLODTQELLQETROKLNLSRRQLEEEK-----NSLQEQEEREARKNQV 1366
QY 253 LDIQALENLEKKEDEILSLKQSLLE-----NIVILSKQVEDLVNVCQLEKEKEDHV 305
Db 1367 LALQSQADTKKKVQDDDDGTTGEEFAKKKLLKDVLEALSQRLEKVLAYDKLETK----- 1422
QY 306 NRNRHNENNAEONMLKQKFLQEQHEKLOQKELOLQDSLQOQBKELSSSHOK----- 360
Db 1423 NRLQELDDLVLVDLH-----QRQIVSNLEKKQKKFDQALLAEKGISARYAERDRAE 1475
QY 361 -----LCSFQEMVKEKNLFEELKQTLDELQKLOQKEQEAERLVKQLEEK 408
Db 1476 AEAREKETKALSARALBEALEAKEEFERQNKQLRADMEDLMSSKDDVGKNVHELEKSR 1535
QY 409 SRAEELKLEKLGKKEAELEKSSNAHTQATLILLQEKYDSMVQSLVEDYTAQFESYKALTA 468
Db 1536 ALBOQVEEMRTQLELEDELEQADAKLRLEVNMQAKAQAFERDLQTRDEQNEEKRLLL 1595
QY 469 SEIEDL--KLENSLSQEKAAKAGK-----NAEDVOHQILATESNOEYVRMLDLQTKSA 521
Db 1596 KQVRELEAELEDERKQRALAVASKKWEIDLKDLQAQTEAANKARDEVIKQLKQAQWK 1655
QY 522 LKTEIKEITVS-----FLQ-----KITDLQNLKQOE-----EDFRKQLED 558
Db 1656 DYQRELEEARSRDEIFAQSKSEKSKKLSLEAEILQLEELASSERARRHAEQERDELAD 1715
QY 559 E-----EGRKA---EK---ENTTAELTEINKWRLLYELLYNKTYPFQOLDADFEVQKA 607
Db 1716 ETANSASGKSALLDEKRLLEARIQALEEELEEEQSNMELLNDRFKTTLOVDTLNTLNTL 1772
QY 608 LLNEHGAQEQNLKTRDSYAKLLGHON--LKOKIKHV-----VKLKDNSOLKSEVSKLR 660
Db 1773 LAASRAQKS-----DNARQOLERQNKELAKQLEEGAVKSKFKATISALEAGIGOLE 1827
QY 661 CQL---AKKQKSETKLOELNKL-----GIKFHDPSPKAFHESKENF-ALKTPLK 707
Db 1828 EQLEQEAERAAANKLVRRTEKKLEIFPMQVEDERRHADQYKEQEKANARMKQLKQLE 1887
QY 708 EGNTCNYRA 716
Db 1888 EAEETATRA 1896

RESULT 7
MLP1_YEAST
ID MLP1_YEAST STANDARD; PRT; 1875 AA.
AC Q02455;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin-like protein MLP1.
GN MLP1 OR YKR095W OR YKR415.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=93247549; PubMed=8483450;
RA Koelling R., Nguyen T., Chen E.Y., Botstein D.;
RT "A new yeast gene with a myosin-like heptad repeat structure.";
RL Mol. Gen. Genet. 237:359-369(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94205265; PubMed=8154186;
RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
RT "The complete sequence of a 15,820 bp segment of Saccharomyces
RT cerevisiae chromosome XI contains the UB12 and MLP1 genes and three
RT new open reading frames.";
```

```

RL Yeast 9:1349-1354(1993).
CC -!- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
CC REPAIR.
CC -!- SIMILARITY: SOME, TO THE TPR ONCOGENE.
CC -!- CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MPL1".
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L01992; AAA34783.1; -
DR EMBL; X73541; CAA51948.1; -
DR EMBL; Z28320; CAA82174.1; -
DR PIR; S38173; S38173.
DR SGD; S0001803; MLP1.
KW Coiled coil; DNA repair.
FT DOMAIN 69 487
FT DOMAIN 531 1678
FT DOMAIN 1834 1866
FT CONFLICT 301 301 R -> A (IN REF. 1).
SQ SEQUENCE 1875 AA; 218455 MW; 683A0D34C9066867 CRC64;

Query Match 11.6%; Score 417; DB 1; Length 1875;
Best Local Similarity 24.1%; Pred. No. 3.8e-07;
Matches 190; Conservative 145; Mismatches 29; Indels 162; Gaps 30;

QY 29 LEVLKGPVSFQSRFKQKQESKQNLNVD---KDTTLPASARKVSKSSSKES--QKND 82
Db 816 LSELAKETS-QKDHHIKQLEED--NNSNIEWQNKTEALKKDYEVITSVDSTQDIEKIQ 873
QY 83 KDLKILEKEI---RVLLQERGAQDRRIQD--LETELEKMEALNAA-----LREKT 128
Db 874 YKVSLEKEIEEDKIRLHTYNYMDETINDDSLKEKSKINLTDAYSQIKYKDIYETT 933
QY 129 SILSANNATLEKQILIELTRTNELLSKFSQNGQ-KNL-----RILSLELMKLN 176
Db 934 SOS-----LQOTNSKLDSEFDFTNQIKNLTDEKYSLEDKISLLEKQMFNLNN 981
QY 177 KRETRMGRMAKQEGMEMKLVQTSRLESQKIAQLEGLKLVSTIEKEIDEKS----- 229
Db 982 ELDLQKQMEKEKADFKKRISILQNNNEVEAVKSEYKSLKIQND-LDQOTIYANTAQ 1040
QY 230 -----ETEKLLYEIEISCASQDQVEKYK-----LDIAQLEENLEKKEKDEILSLKQSL 277
Db 1041 NNYEQELQKHADVSKTISELREQLHTYKGVKTLNLSRDQLENALKENEKSWSSQKESLL 1100
QY 278 ENIVILSKQVEDLVNVCQLEKEKEDHVNRNHNENLNA-----EMQNLKOK 325
Db 1101 EOLDLSNRIEDLSQNLKLLYDIOIYTAADKEVNNTNGPGNLNLTITLRERDILDTK 1160
QY 326 FTLEQEQHEKLOQKELOLQDSLQO-----EKELSSSLHQKLCSPQEE-----MYK 370
Db 1161 YTVAEADAKMLRQKISLMDVDELQDARTKLDNSRVEKENHSSIIQQHDDIMEKLNQLNLLR 1220
QY 371 EKNL-EEEL-----KOTLDELKQKEQEAERLVKQLEEKAKSRAEELKLEKL- 421
Db 1221 ESNITLRNELENNNNKKELQSELQDKLQNVAPIESELTAALKYSMOEKQEQLKAKEEVH 1280
QY 422 ---KGEAELEKSSAAHTQATLILLQEKYDSMVQSL-----DYTAQFESYKALT 467
Db 1281 RWKKRSQDITLEK-----HEOLSSSDYKDESEIENLEKEENKERQGAEEKFNRLRQA 1336
QY 468 ASEIEDLKLENSLSQEK-----AKAGKNAEDVQHQILATESNOEYVRML 513
Db 1337 QERLTKSLQSDSLTEQVNSLRDAKNVLENSLESEANARIEELQNAKVAGNQLEAIRKL 1396
QY 514 L-DLQTKSALKTEIKEITVSFLQKITDILQNLKQOEEDFRKOLEDEEGRKAKEKNTTAE 572
Db 1397 QBDAAKASRELQAKLEESTTSVESTINGLNEETITTLKEIEKQROIQOQQLQATSANEQND 1456
```



Db 1773 LAERSAQAQS-----DNARQQLERQNKELKAKIOLEAGAVKSKFKATISALEAKIGOLE 1827  
 Qy 661 CQL---AKKQSETKLOELNKVL-----GTXHPDPSKAPHESKENF-ALKTPLK 707  
 Db 1828 EQLQEAKERAANKLVRRTEKKLEIFEMQVEDERRHADQYKEQMEKANARKQLKROLE 1887  
 Qy 708 EGNVNCYRA 716  
 Db 1888 EAEEEAATRA 1896

RESULT 9  
 MYS2\_DICDI STANDARD; PRT; 2116 AA.  
 AC P08799;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin II heavy chain, non muscle.  
 GN MHCA.  
 OS Dictyostelium discoideum (slime mold).  
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87092266; PubMed=3540939;  
 RA Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;  
 RT "Conserved protein domains in a myosin heavy chain gene from  
 Dictyostelium discoideum";  
 RL Dictyostelium discoideum";  
 RN Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).  
 RP [2]  
 RP PHOSPHORYLATION SITES, AND MUTAGENESIS.  
 RC STRAIN=AX2;  
 RX MEDLINE=90353583; PubMed=2387408;  
 RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,  
 RA Gerisch G.;  
 RT "Replacement of threonine residues by serine and alanine in a  
 phosphorylatable heavy chain fragment of Dictyostelium myosin II";  
 RL FEBS Lett. 269:239-243(1990).  
 RN [3]  
 RP PHOSPHORYLATION SITES.  
 RX MEDLINE=88112226; PubMed=2828113;  
 RA Magle G., Noegel A., Scheel J., Gerisch G.;  
 RT "Phosphorylation of threonine residues on cloned fragments of the  
 Dictyostelium myosin heavy chain";  
 RL FEBS Lett. 227:71-75(1988).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.  
 RX MEDLINE=95345066; PubMed=7619795;  
 RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,  
 RA Rayment I.;  
 RT "X-ray structures of the myosin motor domain of Dictyostelium  
 discoideum complexed with MgADP.Befx and MgADP.ALf4-";  
 RL Biochemistry 34:8960-8972(1995).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.  
 RX MEDLINE=95345067; PubMed=7619796;  
 RA Smith C.A., Rayment I.;  
 RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the  
 truncated head of Dictyostelium discoideum myosin to 2.7-A  
 resolution";  
 RL Biochemistry 34:8973-8981(1995).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
 RX MEDLINE=96206189; PubMed=8611530;  
 RA Smith C.A., Rayment I.;  
 RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the  
 Dictyostelium discoideum myosin motor domain to 1.9-A resolution";  
 RL Biochemistry 35:5404-5417(1996).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.  
 RX MEDLINE=97452580; PubMed=9305951;  
 RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;

RT "X-ray structures of the MgADP, MgATPgammAs, and MgAMPPNP complexes  
 of the Dictyostelium discoideum myosin motor domain";  
 RL Biochemistry 36:11619-11628(1997).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
 RX MEDLINE=98070605; PubMed=9405148;  
 RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;  
 RT "X-ray crystal structure and solution fluorescence characterization  
 of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the  
 Dictyostelium discoideum myosin motor domain";  
 RL J. Mol. Biol. 274:394-407(1997).  
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE  
 ACTIVITY THAT IS ACTIVATED BY ACTIN.  
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES  
 INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI  
 LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
 (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL  
 CORTEX.  
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER  
 SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 SUBFRAGMENT (S2).  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES  
 THE ACTIN-ACTIVATED ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE  
 ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1  
 POSITION (688).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; M14628; AAA33227.1; -;  
 PIR; A26655; A26655.  
 PIR; S00250; S00250.  
 PDB; 1MMA; 03-DEC-97.  
 PDB; 1MMD; 17-AUG-96.  
 PDB; 1MMG; 03-DEC-97.  
 PDB; 1MMN; 03-DEC-97.  
 PDB; 1MND; 17-AUG-96.  
 PDB; 1MNE; 17-AUG-96.  
 PDB; 1VOM; 23-DEC-96.  
 PDB; 1LVK; 28-JAN-98.  
 DictyDb; DD01008; mhca.  
 InterPro; IPR000048; IQ region.  
 InterPro; IPR004009; Myosin\_N.  
 InterPro; IPR001609; myosin\_head.  
 Pfam; PF00063; myosin\_head; 1.  
 Pfam; PF00612; IQ; 2.  
 Pfam; PF02736; Myosin\_N; 1.  
 PRINTS; PR00193; MYOSINHEAVY.  
 ProDom; PD000355; myosin\_head; 1.  
 SMART; SM00015; IQ; 1.  
 SMART; SM00242; MYSC; 1.  
 PROSITE; PS50096; IQ; 1.  
 Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;  
 Calmodulin-binding; Methylation; Alkylation; Phosphorylation.  
 FT DOMAIN 1 761 MYOSIN HEAD-LIKE.  
 FT DOMAIN 762 791 IQ.  
 FT DOMAIN 817 2116 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP.  
 FT DOMAIN 638 660 ACTIN-BINDING.  
 FT DOMAIN 738 752 ACTIN-BINDING.



```
FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).
FT MOD_RES 678 678 ALKYLATION (SH-1).
FT MOD_RES 1823 1823 PHOSPHORYLATION (BY MHCK).
FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MHCK).
FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BB1EE56AI CRC64;

Query Match 11.5%; Score 414; DB 1; Length 2116;
Best Local Similarity 22.3%; Pred. No. 5.4e-07;
Matches 191; Conservative 178; Mismatches 275; Indels 214; Gaps 36;

QY 8 LKRFNDPSCAIPGAYD-VKYLEVLKGPVSPQK-----SQRFKQKQSKQNLNVDK---- 58
DB 959 MKRVND-----GQSDIRISLEIKD--ELQKEVEELTESFSESKDKGVLEKTRVRL 1008
QY 59 -----DTTLPASARKVKSESKQKNDKILKEITRVLLQERGAQDRIOQLETEL 113
DB 1009 QSELDLTVRLOSEYKDKSELRLQKKLEELKQVEALAEATAKLAQEAANKKLQGEY 1068
QY 114 EKMEARLNAALREKTSLSANNATLEKQLIELTRTNELLSKSFSENGNOKNRLRILSLELMK 173
DB 1069 TELNEKFNESEVARSNVESKKTLESQLVAV--NNELDEEKNRDALEKKKKALDAMLLEE 1126
QY 174 LRNKRETKRGMMAQEGEMK-----LQVQRS-LEESQKTAQLEGKLVSTEEKIDBK 228
DB 1127 MKDQLES-----TGGEKSLYDLKVKQESDMEALRNOISELQSTIAKLEIKKSTLE 1177
QY 229 SETEKLEVEIEISCASDOVEKYK-----LDI-----AQLEENLEKKNDETLSLKOSLEENI 280
DB 1178 GEVARLOGLEAEQAQAKSNVEKQKKVELDLEDKSAQAEATAK-QALDKLKKLEQEL 1236
QY 281 VILSKQVEDLVNKKCOLLEKEKEHDVNRNREHNENLNAEQNLKQFLEQKQHEKLOOKE 340
DB 1237 SEVQITQLSAANK-----NVNSD-STNKHLETSPNNLKLEAEQAKQALEKKR 1285
QY 341 LAQDS-----LLOQEKLSLHQLKCSFQEMVKENLFEELKQTLDELDKLOOKE 393
DB 1286 LGLESELKHVNQLEEFKQKESKRRKVDLEKVEYSELQDQIEEVASKVAYTEAKRKE 1345
QY 394 EQAERLVQLEEAESR-----AEELKLEK-----LKGKAELEKSSAAHTQATLLQ 443
DB 1346 SELDEIKROYADVSRDQKSVQLTKLQAKNEELRNTAEABQGLDRAERSKKAEFDLE 1405
QY 444 EKYDSWQSLDVTA-QFESYKAL-----TASEIEDLK----- 475
DB 1406 E-----AVKNLEETAKKVAEKAMKAETDYRSTRKSELDDAKMVSSEQVYQIKRLNEELS 1461
QY 476 -----LE-----NSLQEK-----AAKAGKNAEDVOHOILATES 504
DB 1462 ELRSVLEEADERCNSAIKAKKTAESALESKDEIDAANNAKAPRKSKELEVRVAELFE 1521
QY 505 S-----NQEVV-----RMLLDLQTKSALKETEIKE----- 529
DB 1522 SLEDSKSGTVNVEFIRKDAEIDDLRLDREFESIKSDEDEKKNTRKQFADLEAKVERAQ 1581
QY 530 -----ITVSFLQK-----ITDLQNLQKQE-----EDFRKQLEDE-EGRKAERKNTTAEIT 574
DB 1582 REVVTIDRLKKKLESIDIILSTQLOTETKSRKIEKSKKKLQTLAERRAEEGSSKAAD 1641
QY 575 EEINKWRLLYEELYNKTRFPQLODAFEVEKQALLNEHGAQEQNLKTRDSYAKLLGHON 634
DB 1642 EIRK-----QVWQEVDELRAQLDS-----ERAALN-----ASEK-KIKSLVAEY---DE 1682
QY 635 LKQKTKHVVKLDE-----NSQLKSEVSKLRCOLAKKQKSETKLOEELNKVLGIKHFDPSK 690
DB 1683 VREKLEIDEILAKDLVKAKRALEVELEEVROQLEEEEDSRSEL-EDSKRRLLTTEVEDIKK 1741
QY 691 APFHESKENFALKTPLE 708
DB 1742 KYDAEVEQNTKLEAKK 1759
```

RESULT 10

```
MYH4_HUMAN
ID QYV623; STANDARD; PRT; 1939 AA.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, fetal (Myosin heavy chain IIB)
DE (MYHC-IIB).
GN MYH4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=93318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CHARACTER OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF111783; AAD29949.1; -
DR HSP; P13538; 2MYS.
DR Genew; HGNC:7574; MYH4.
DR MIN; 160742; -
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF01576; Myosin_tail; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
FT DOMAIN 785 814 IQ.
FT DOMAIN 843 1939 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT SEQUENCE 1939 AA; 223012 MW; 40BLADID777A47DE CRC64;

Query Match 11.4%; Score 409.5; DB 1; Length 1939;
```



```
Best Local Similarity 22.7%; Pred. No. 6.9e-07;
Matches 186; Conservative 131; Mismatches 233; Indels 271; Gaps 30;

QY 136 TLEKQLLE-----LTRNELLSKFSFENGQKNLRILSLMLRNKREY-----180
Db 776 TLEMRDEKLAQLTRFOALCRG-----FLMRFEKRMERRESFICQYINRA 824
QY 181 -----KMRGNMAKOE-----GMEMKQVQORSLESQKIAQLEGKLVSI 220
Db 825 FMNVKHPWPKLYKIKPLPKSAETEKEMANKEEFKTEELAKTEAKKELEEKVYL 884
QY 221 EKEKIDSKSTKLELVIEISCASDOVEYKYL-----DIAQLENLKEKNDIEILSK 273
Db 885 MOEKNDLQVOAEADALADAEERCDOLIKTQLEAKIKVEYARDEEINAEITAKK 944
QY 274 QSLLENIVILSKOVEDLVNVCOLLEKEDHVRNRNHNENLNAEQNL-----KOK 325
Db 945 RKLDECESELKDDLDLFLAKVEKEK--HATENKV--NRLTEEMAGLDETAKLTKR 1000
QY 326 FILE---QOEHEKIQKELQIDSL-----QOEKELSSSLHQ--KLCFSQE-----366
Db 1001 KALQEAHQQTLDLQMEEDKVNTLTAKTKLEQQVDDLEGSLEQEKLCMDLERAKRKE 1060
QY 367 ---EMVKEKNLFEELKQTDLDKLOOKE-----EOA-----396
Db 1061 GDLKLAQESTMTDENDKQOLNE--KLKKKEFEMSNLQKIEDEQALAMQLKKIKELQAR 1118
QY 397 -ERLVKOLEEAKSRAB-----ELKLEKEL-----KGRKAE-----427
Db 1119 IEELFEIEERASRAKAEKQSDLSRELEISERLEEAGGATSAQIELNKKREAEFQKM 1178
QY 428 ---LEKSSAAHTQATLLQEKY-DSM-----VQSLDVTQAQFESYKALTASETEDLKE 477
Db 1179 RRDLEESTLQHEATAAALRKKHADSVAEKQIDSLQVRVKQKLEKESLKMELNDLASN 1238
QY 478 NSSLQEAAKAGNAEDVQHIILATESNQYVRLMLDLOTKSALKETETKEIT-----531
Db 1239 METVSKAKANFECMRTLEDQSLSEIKTEEBEQORLINELSAQKARLHTESEGFSQLDEK 1298
QY 532 -----VSFLQKITDLOLQKOEED-----FRKQLEDEGRK 563
Db 1299 DAMVSQLSRQKQATQOIEELKQLEETKAKSTLAHALASARHDCDLLREQVEEEOEAK 1358
QY 564 AEKENTTAELTEETINKRWLLY-----BELYNKTRPFQLOLD-----AF 601
Db 1359 AELQRGMSKANSEVAQWRTYETDAIQRTTELEAKKLAQRLQDAEEHVEAVNSKASL 1418
QY 602 EVEKQALLNE-----HGAQEQNLKIRDSYAKLLGHQNLK-----636
Db 1419 EKTQRLQNEVEDLMDIDVERSNAACIALDKKQRFNFKVLAEWKQKYEETQAELEASQKES 1478
QY 637 -----QKIKHVVKLDENSOLKSPVSKRLCOLAKKKQSETKLOEELNKVL 681
Db 1479 RSLSTELFKVNAYEESLDHLETKRNKNLQQISDLTQIAE-----GGKHIELEKV- 1533
QY 682 GIKHFDPSKAFHHESKENFALKTPKLEGNTNCRAPMECOE 722
Db 1534 -----KKQLDHEKSE---LQTSLEEA-----EASLEHEE 1559

RESULT 11
MYH9_CHICK
ID MYH9_CHICK STANDARD; PRT; 1959 AA.
AC P14105;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MYosin heavy chain, nonmuscle (Cellular myosin heavy chain) (NMHC).
GN MYH9.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
```

```
NCBI_TaxID=9031;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Intestinal epithelium;
RX MEDLINE=90046668; PubMed=2813355;
RA Shohet R.V., Conti M.A., Kawamoto S., Preston Y.A., Brill D.A.,
RT Adelstein R.S.;
RT "Cloning of the cDNA encoding the myosin heavy chain of a vertebrate
cellular myosin.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7726-7730(1989)
CC -!- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CAPPING.
CC -!- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC-2),
REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
EMBL; M26510; AAA48974.1; -.
DR PIR; A33977; A33977.
DR HSSP; P10587; 1BR2.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Alkylation; Multigene family.
FT DOMAIN 1 778 MYOSIN HEAD-LIKE.
FT DOMAIN 779 808 IQ.
FT DOMAIN 837 1925 COILED COIL (POTENTIAL).
FT NP_BIND 174 181 ATP (POTENTIAL).
FT DOMAIN 654 676 ACTIN-BINDING.
FT MOD_RES 694 694 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
SQ SEQUENCE 1959 AA; 226502 MW; A75C86086FD3A1A1 CRC64;

Query Match 11.4%; Score 409; DB 1; Length 1959;
Best Local Similarity 23.0%; Pred. No. 7.2e-07;
Matches 189; Conservative 159; Mismatches 289; Indels 184; Gaps 32;

QY 37 SFQKSQRFKQKQKQNLVNDKDTTLTPASARKVKVSSSKESQK--NDKDLKILEKEIRV 94
Db 873 TFO-AQLMAEKMQLOEQLOAEALCAEAEEIRAKLTAKKQLELEICHDLARVEEERC 931
QY 95 --LLQERGAQDRRIQDLETELEKME-ARLNAALREKTS-----LSANNATLEKOLIELT 145
Db 932 QHLQAEKKKMQQNTQEELEEEESARQKLEKVTTEAKLKKLEEDVIVLEQNLKLA 991
QY 146 RTNELLSKSESE-NGNOKNRLRILSLMLKLNKRE---TKMRGMMAKQEGMEKMQVTOR 201
Db 992 KEKKLLEDRWSEFTNLTTEEEKSKSLAKLNKHEAMITDLEERLRREKQORQELEKTR 1051
```

```

QY 202 SLE-----ESQKIA-----OLEKGLVSIKE-----KIDE-KS 229
D 1052 KLEGGSDLDHQIAELQIAELKIQLSKEEELAAALAEVEEAQKNMALKRIELES 1111
QY 230 ETEKLEYIEEISCASQDQVEKVKLDIAOLEENLKEKNDIEL-----SLKOSLEENIVI 282
D 1112 QITEQOEDLESASRKNKAQKRDGLGEELEALKTELEDTDSTAQQOELRSKRQEYTV 1171
QY 283 LSKQVED-----LNVKQQLLEKEKEDHVN 306
D 1172 LAKTLEDEAKTHEAQIOENRQKHSQAIBELAQLEOTKRVKANLEKAKQALESERAEISLN 1231
QY 307 R-----NREH-NENLNAEQMLKQKFI-----LEQOEHEKLOQKELOIDS-----LLQQ 349
D 1232 EVKVLLOGKDAEHRKKVYDAQLOELQVKTGERVKTLEAERVNKLQVELDQNVGTLNQ 1291
QY 350 EKELS-----SSLHOKLSQOEWEKKNL---FEELKQTLDELKLOQKEQOERL 399
D 1292 SDSKSIKIAKDFSALESQIQDTQELLQOETRLKLSFSTKLQTEDEKNALKBQLEEEEA 1351
QY 400 VKOLEE-----EAKSRAEE---LKLLEELKGEAELEKSSAAHTQ---ATLLL 442
D 1352 KNLEKQISVLQQAQVAREKMKDDGLGCLIEAEBAKKLQKLDLESITORYEBKIAAYDKL 1411
QY 443 QEKYSMVQSLDVTAFESYKALTASETE-----DLKLENSLOEKAA-----KAGKN 491
D 1412 ETKTKRLQQLDLDIAVDLD-HQQTQVSNLEKKQKQKFDQLLAEKNISAKYAEERDRAEA 1470
QY 492 AEDVQHOILATSSNQEVYRMLDQTSALKETIKEITVS-----FLOKITDLOQL 545
D 1471 AREKETKALSTARALEAEAEQKAELEVRNKPQRTMEDLMSSKDDVGVSHLEKAKRAL 1530
QY 546 KOEEDEFKQLEDEGRKAERKNTTAELETEINKWR-----LLEELYNKTKPFOL- 596
D 1531 EQOVEMTKQLELEDELOATED--AKLLEVNQQAQKAFDRDLGDRDEQNEERKARKLI 1588
QY 597 -QIDAFQEKQALLNEHGAQSQLNKRIRDSYAKLIGH-----QNLKQKIKHVKLKDENS 650
D 1589 RQVREMEVELEDERKQSRSTAVAAKRLDLDKLESHDITANKNRDEAIKHVKRLQAQMK 1648
QY 651 QLKSEVSKLRQ----LAKKKQSETK-----LQEL 677
D 1649 DYMRELEDTRTSREBILAQAKENKRLKLSMEAEIMQLQEL 1689

RESULT 12
MYSN_DROME
AC Q99323;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, non-muscle (zipper protein) (Myosin II).
GN ZIP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=90349606; PubMed=21177279;
RA Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;
RT "Complete sequence of the Drosophila nonmuscle myosin heavy-chain
RT transcript: conserved sequences in the myosin tail and differential
RT splicing in the 5' untranslated sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
CC -!- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
CC CELLULARIZATION REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

```

```

CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M35012; AAA28713.1; -.
CC PIR; A36014; A36014.
CC PIR; B36014; B36014.
CC HSSP; P10587; 1BR2.
CC FlyBase; FBgn005634; zip.
CC InterPro; IPR000048; zip.
CC InterPro; IPR000049; IQ_region.
CC InterPro; IPR004009; Myosin_N.
CC InterPro; IPR002928; Myosin_tail.
CC InterPro; IPR001609; myosin_head.
CC Pfam; PF00063; myosin_head; 1.
CC Pfam; PF00612; IQ; 1.
CC Pfam; PF01576; Myosin_tail; 1.
CC Pfam; PF02736; Myosin_N; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC PRODOM; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MYSC; 1.
CC PROSITE; PS00096; IQ; 1.
CC Myosin; Alternative splicing; Coiled coil; Actin-binding;
CC ATP-binding; Calmodulin-binding.
CC FT DOMAIN 1 829
CC FT DOMAIN 830 859
CC FT DOMAIN 886 2017
CC FT NP_BIND 225 232
CC FT DOMAIN 250 260
CC FT DOMAIN 682 694
CC FT DOMAIN 705 727
CC FT DOMAIN 742 758
CC FT DOMAIN 1303 2017
CC FT DOMAIN 1303 1970
CC FT DOMAIN 1971 2017
CC FT VARSPLIC 1 45
CC SQ SEQUENCE 2017 AA; 232016 MW; 73E3CB02BA8F2528 CRC64;

Query Match 11.3%; Score 408; DB 1; Length 2017;
Best Local Similarity 24.7%; Pred. No. 8e-07;
Matches 209; Conservative 134; Mismatches 286; Indels 218; Gaps 35;

QY 29 LEVLKGPVSF-QKSQRFKQKESKQNLN-----VDKDT-----TLPA 64
D 889 LEVTKQBEKLVQKEDELKQVREKLDTLAKNTQEVERKYQQALVEKTTLAEQLQAEIELCA 948
QY 65 SARVKVS-----SESKKESQKNDKDLKILEKIRVLL--QERGAODRRIOQDLELEKWEA 118
D 949 EAEESRSLMARKQELDMQOELETRIEEERVLALGEGKKLELNQDLEQLEEEEA 1008
QY 119 RLNAALREKTSLSA-----NNATL-EKOLIELTRTNELLKSKFSFENGKQ 162
D 1009 ARQKLQLEKQVOLDAKIKKYEDLALTDQNKLLKEKKLLB-ERANDLSQTLAEEEKAK 1067
QY 163 NLRILSLKMLKRNKRE---TKMRGMMAKQEGMEMKLVQVQTSRLE-----ESQG 208
D 1068 H-----LAKLAKHEATITELERLHKDQOQOQESDRSKRKIETEVDLKEQLNERRV 1120
QY 209 KIAOLEKGLVSIKE-----KIDEKETE-----KLLEYIEEISCASDQ 247
D 1121 QVDEMQAQLAKREBELQTLLRIDEESATKATAQAQKQRELSQAEIQEDLEAEKAAKAK 1180
QY 248 VEKVKLDTAOLEENLK-----EKNDLEILSKOSLEENIV-----281
D 1181 AEKVRDLSELEALKNELLDLSLTTAAQQLRSKRQELATLAKSLEETVNHGVLAD 1240
QY 282 -----ILSKQVEDLNVKQCLEKEKED-----HVNRRNHNENLN-- 316

```



QY 430 KSSAHTQA-----TLLQKQYDSWQSLDVT-----QFESYKALTAETE 472  
Db 821 KNLAVNQKDSLEKELQKKEFTSAVDAENARQAMQETINKLNKEEQFALMSSELE 880  
QY 473 DLKLENSLOEKAARAKAGNAED-----VOHQILATESSNOEYVRLMLDLOTKSAKE 524  
Db 881 QLK-SNLTYMETKLEREREQOQLTEAKYKLENDIAETIMKSSGSSAQLMKMNDLRLKE 939  
QY 525 TEKETVTSFL--OKITDLOLQKQ-----QEE-----DFRKOL 556  
Db 940 RLEQIQLELTAVANEKAVOLQKNVQTAQKABQSOQETLKHQEEELKMKQDLOTKMKOM 999  
QY 557 EDEEGR---KAENKENTAELEEN-----KWRLLYBELYNKT---KPFOLQLDAPFV 603  
Db 1000 ETSQNYKDLQKYEKETSEMITKHADIKGFKQNLDAEALKAQKKNDELETAQEEEL 1059  
QY 604 EKQALLNEHGAQEQNLKTRDSYAKLLGHQNL--KQKTHVVKLKDENSQKSEVSKURC 661  
Db 1060 KKOAA---EQAKADKRAEEVLOTMERVTRKDAIHOEKIETTLASL--ENSROTN--KLQN 1112  
QY 662 OLAKKKQSEKLOEELNVLGKTHFDPSKAFPHESKENF-ALK 703  
Db 1113 ELDMLKQNNKNEBELTSKELLNENKV--BELKKEFEALK 1153

RESULT 14  
MYH2\_HUMAN  
ID MYH2\_HUMAN STANDARD; PRT: 1941 AA.  
AC Q9UKX2: Q16229; Q14322;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin heavy chain, skeletal muscle, adult 2 (Myosin heavy chain IIA)  
DE (MyHC-IIa).  
GN MYH2 OR MYHSA2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=99318869; PubMed=10388558;  
RA Weiss A., Schiaffino S., Leinwand L.A.;  
RT "Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity";  
RL J. Mol. Biol. 290:61-75(1999).  
RN [2]  
RP SEQUENCE OF 1711-1941 FROM N.A.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=95109625; PubMed=7545970;  
RA Smerdu V., Karsch-Mizrachi I., Campione M., Leinwand L.,  
RA Schiaffino S.;  
RT "Type Iix myosin heavy chain transcripts are expressed in type IIB fibers of human skeletal muscle";  
RL Am. J. Physiol. 267:C1723-C1728(1994).  
RN [3]  
RP SEQUENCE OF 1823-1941 FROM N.A.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=95270723; PubMed=7751403;  
RA Enlioni S., Sant'Ana Pereira J., Sargeant T., Young A., Goldspink G.;  
RT "Characterization of human skeletal muscle fibres according to the myosin heavy chains they express";  
RL J. Muscle Res. Cell Motil. 16:35-43(1995).  
CC -1- FUNCTION: MUSCLE CONTRACTION.  
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEATITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY

CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
CC EMBL; AF111784; AAC29950.1; -.  
CC EMBL; S73840; AAC13916.1; -.  
CC EMBL; P23258; CAAB3687.1; -.  
CC HSSP; P13538; 2MVS.  
CC Genew; HGNC:7572; MYH2.  
CC MIM; 160740; -.  
CC InterPro; IPR000048; IQ\_region.  
CC InterPro; IPR004009; Myosin\_N.  
CC InterPro; IPR002928; Myosin\_tail.  
CC InterPro; IPR001609; myosin\_head.  
CC Pfam; PF00063; myosin\_head; 1.  
CC Pfam; PF00612; IQ; 2.  
CC Pfam; PF01576; Myosin\_tail; 1.  
CC Pfam; PF02736; Myosin\_N; 1.  
CC PRINTS; PR00193; MYOSINHEAVY.  
CC ProDom; PD000355; myosin\_head; 1.  
CC SMART; SM00015; IQ; 1.  
CC SMART; SM00242; MYSC; 1.  
CC PROSITE; PS50096; IQ; 1.  
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
KW Multigene family.  
FT DOMAIN 1 786 MYOSIN HEAD-LIKE.  
FT DOMAIN 787 816 IQ.  
FT DOMAIN 845 1941 COILED COIL (POTENTIAL).  
FT NP\_BIND 179 186 ATP (POTENTIAL).  
FT CONFLICT 1844 1844 K -> R (IN REF. 3).  
SQ SEQUENCE 1941 AA; 223043 MW; 681E866F83AE83F CRC64;  
Query Match 11.38; Score 407; DB 1; Length 1941;  
Best Local Similarity 24.08; Pred. No. 8.3e-07;  
Matches 211; Conservative 136; Mismatches 253; Indels 278; Gaps 39;  
QY 51 KQNLNVKDTTLPASARKVSKESKQKNDKDLKLEKIRV-----LLQ 97  
Db 725 KQRYVLNASEIPEG---QFIDSKKASEKLLASIDIDHTQYKFGHTKVFVKAGLLGLE 780  
QY 98 ERGAQDRRIQDLETELEKMEARLNALREKTSLSANNATLEKQIETLTRNELLKSKFSE 157  
Db 781 E--NRDDKLAQLIT---RTQARCRGFLAR-----VEYQRMVERREAFICI 820  
QY 158 NGNKN-LRILSLELMKLRNRETMRGMMAKQSGEMK--LQVTSRLESQKIAOLE 214  
Db 821 QYNTRSFNNVHWPMWPKLFFKIPKLKSAETEKEMATMKKEEFQIKDELAKSEAKRKELE 880  
QY 215 GLKVSIEKIDEKSETEKLELEYIEEISCASDQVEKYKL-----DIAQLEENLKEKND 267  
Db 881 EKMYTLKKNKNDLQVQAEAGLADAEERCQDLIKTIQLEAKIKVEYTERADEEEINA 940  
QY 268 EILSKOSLENIIVLSKQVEDLVNKVQOLLEKEKEDHVRNREHNENUNAMQNL----- 322  
Db 941 ELTAKRKLDECELSKKDIDDLTLLTAKVKEK--HATENKV--KNITEMAGLDEITIA 996  
QY 323 ---KQKFTLE-----QOEHEKIQ-----QKLEQIDSL---LQOEKELSSSLHQ 359  
Db 997 KLTRKALQEAHQHTLDDLDLQAEEDKVNLTAKIKLEQQVDDLEGSLEQEKLLMDLER 1056

QY 360 -----KLCFSQEMV---KEKNLFEELKQTLDELKLOQK---EEQA----- 396  
Db 1057 AKRKLEGLKLA--QESIMDIENEKQOLDEKLLKKEFEISNLQSKTEDEQALGIQKKI 1114  
QY 397 -----ERLVKOLEPEAKSRAE-----ELKLLBEKL-----KGRE 425  
Db 1115 KELQARIEELEIEAEASRAKAEKORSDSLSELEIEISERLEBEAGGATSAQTEMNKKRE 1174  
QY 426 AE-----LEKSSAHTQATLLQEKY--DSM-----VQSLBDVTAQFESYKALTASEI 471  
Db 1175 AEFQKMRDLLEATLQHEATAATLRKKHADSVAEELGEQIDNLRQVKQKLEKESMKMEI 1234  
QY 472 EDLKLNESSLOEAKAKAGNAE----DVOHQILATESSNQEYVRMLLD-----LQTKS 520  
Db 1235 DDL-----ASNVTVSRAKGNLEKMCRTLEDQLSLKSEKEEQORLINDLTAQRGRLOTES 1290  
QY 521 A--LKETEKEIYVS-----FLQKITDQLNQLKQO-----EED---FR 553  
Db 1291 GEFSSRLDEKALVSQSRGKQAFQTOOIBELKQOLEEIEKAKNALAHALQSSRHDCDLLR 1350  
QY 554 KQLEDEEGRAKAKENTTAELTEIEINKWRLLY-----EELYNKTKPFQQLDA----- 600  
Db 1351 EQYEEESKAEALQRLALSKANTEVAQWRKYETDARTQTEEEEAANKLAORLOAAEEHV 1410  
QY 601 -----FEVEKQALLNE-----HGAAQPOLNKIRDSYAKLLG-----H 632  
Db 1411 EAVNAKASLEKTKQRLQNEVEDLMDLMDVNTNAACAALDKKORNFDKILAEMKQKCEETH 1470  
QY 633 QNLKQ-----KIKHVVK-----LKDENSOLKSEVSKRLCOLAK----- 665  
Db 1471 AELEASQKARSGLTELFKIKNAYESLQDLETLRKNKNLQOEISDLTEQIAEGGKRIH 1530  
QY 666 ----KKQ-----SETKLOEELNKVLGK 684  
Db 1531 ELEKIKQVEQEKCELOQAALAEAEASLEHEEGKILRIQ 1568

RESULT 15  
Y109 YEAST  
ID Y109 YEAST STANDARD; PRT; 1679 AA.  
AC P40457;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Hypothetical 195.1 kDa protein in DNA43-UBI1 intergenic region.  
GN Y1149C  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
RA Gentles S., Hamlin N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
RA Louis E., Lye G., Mould S., Mould T., Odell C., Pearson D.,  
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
RA Walsh S.V., Whitehead S.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z38059; CAA86129.1; --  
DR F1R; S48385; S48385.  
DR SGD; S0001411; MLP2.  
KW Hypothetical protein.

SQL SEQUENCE 1679 AA; 195141 MW; 298950CC52202D8F CRC64;  
Query Match 11.3%; Score 405; DB 1: Length 1679;  
Best Local Similarity 21.9%; Pred. No. 8.5e-07;  
Matches 200; Conservative 167; Mismatches 321; Indels 224; Gaps 32;  
QY 25 DVKTLEVLKGPVGFQSRFQKQKESQKQNLNVDKDTTLPASARKVKSSESQKNDKD 84  
Db 471 NTSAIQETASPLSQDELISLRKILESSNIVN-ENDSQAIITERLVFESNVNELQKNVEL 529  
QY 85 LKILEKEIRVLQOE---RCQADRRITODLET-----ELE-----KWEARLNAALR 125  
Db 530 LNC-----IRIADKLENYEGKDTLQKVENQTIKEAKDAIIELENINAKMETRINILLR 585  
QY 126 EK-----TSLSANNATLEKOLIELTRTNELLKSKSPENGOKNLRILSL 169  
Db 586 ERDSYKLLASTEENKANTNSVTWEAAREKKIRELEA--ELSTKYVENSATIONLRKELL 643  
QY 170 ELMKLRNKRETKM-----RGMMAKQEGM-----EMKLVQTOR----- 201  
Db 644 IYKKSQCKKTKTTLEDPENFKGLAKERMLLEAIDHLKAELEKQKSWPSYIVHVEKERAS 703  
QY 202 -SLEESQGKTAQLEGLVSTIEKEKID-----EKSTELKLEYIEIESC 243  
Db 704 TELSOSRIKIKISLEYEISKLKKEKTASPTKESLTRDFEQCCKEKKELQMLRKE-SEISH 762  
QY 244 ASDQVE-----KYKLDIAOLEN-----LKEKNDEILSKOSLEENIVILSKOVEDL 290  
Db 763 NENKMPFSKEGQTKAKIKELNENLERLSDLQSKIOEISIRSCDSQKLNQANTIDDT 822  
QY 291 NVKCOLLEK---KEDHVNRRNHNENLNAEMNLK--QKFI-----LEQOEHEKLQ 337  
Db 823 EMKMSLLTSLNKEITIEKLSSEIENLDKELRKTQYKFLQDQNSDASTLEPTLRKELE 882  
QY 338 QKELQI---DSLLOQEKELSSHLQKLCFQOEEMVKKNLP-----EELK 380  
Db 883 QIQVOLKDANSQIQAYEEIISSENALIELKNELAKTKENYDAKIELEKKEKWAREEDLS 942  
QY 381 QTLDEL-----DKLQKEEQEARLVKOLE----- 404  
Db 943 RLKELGETRALQPKLKEGALHFVQOSEKURNEVERQKMKIEKEMSTIVQLCKKKEMS 1002  
QY 405 -----EAKSRAEELKLEELKLEKAELEKSSAAHTQATLLQEKYDSVMVQSLQEDVTA 458  
Db 1003 QYOSTMKENKDLSELVIRLEKDAADCAQELTKTKSSLYSA---QDLDRKHERKWEKA 1058  
QY 459 QFESYKALTASEITEDUKLNESSLOEAKAKAGNAEDVOHOILATESSNQEYVRMLLDIQT 518  
Db 1059 DYERELISNTEQTESLRVENSVLIEKVDDTAANNNGDKDHLKLVSLFSLNLRHNSLETKL 1118  
QY 519 KSALKETE-IKEITVSPLOKITDLOLQKQEEDEF--KQLEDEGRKAEKENTAEELTEE 576  
Db 1119 TTCRELAFYKQKNDLSLEKTIINDLQRTQTUSEKEYQCSAVIIDFKDITKEVTOVNILKE 1178  
QY 577 IN-----KWRLLYEELYNKTKPF--QLQDAFEVEKQALLNEHG--AAEQOLNK 621  
Db 1179 NNAILQKSLKNVTEKNEIYKQNDROEISRLQORDLIQTEQVINSKILYVESEMEQ 1238  
QY 622 IRDSYAKLLGHQNLKQKIKHVVKLKDENSOLKSEVS-----KLRQLAKK- 666  
Db 1239 CKQRYQDLSQQQDAQK-KDIEKLTNEISDLKGLKLSAENANADLENKFNRLKQKAEKL 1297  
QY 667 ---KQSETKLOEELNKVLGK-----HFDPSKAFHESK---ENFALTKPLKEGWTNC 713  
Db 1298 DASKKQQAATNELNELKAIKDKLEQDLHFENAKVILDLDTKLKAHELQSEDSVNDHEKDT 1357  
QY 714 YRAPMECOESWK 725  
Db 1358 YRTLMEEIESLK 1369



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:09:23 ; Search time 19.7824 Seconds  
(without alignments)  
3523.197 Million cell updates/sec

Title: US-09-685-010-47

Perfect score: 3598

Sequence: 1 MSFPKAPLKRFPNDPSCGAPS.....LKEGNTNCRAPMECQESWK 725

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3598	100.0	725	1 JC5016	hyaluronan recepto
2	1952	54.3	631	2 JC4298	hyaluronan recepto
3	444	12.3	1790	2 S67593	transport protein
4	442	12.3	2139	2 T18296	myosin heavy chain
5	436.5	12.1	1690	2 T13030	microtubule bindin
6	425.5	11.8	1410	1 A57013	early endosome ant
7	417	11.6	1875	2 S38173	myosin-like protei
8	416	11.6	1738	2 T14867	interaptin - slime
9	414	11.5	1776	2 A59252	myosin heavy chain
10	414	11.5	2116	2 A26655	myosin heavy chain
11	409	11.4	1959	1 A33977	myosin heavy chain
12	409	11.4	2007	1 B43402	myosin heavy chain
13	408	11.3	2017	1 A36014	myosin heavy chain
14	408	11.3	2057	2 S61477	myosin II heavy ch
15	405.5	11.3	1964	2 A59282	nonmuscle myosin I
16	405	11.3	1679	2 S48385	hypothetical prote
17	405	11.3	1961	1 A61231	myosin heavy chain
18	403	11.2	1938	2 A59293	skeletal myosin he
19	403	11.2	1940	1 A24922	myosin heavy chain
20	402.5	11.2	1972	1 A41604	myosin heavy chain
21	402	11.2	1979	1 S03166	myosin heavy chain
22	400	11.1	1999	1 S21801	myosin heavy chain
23	399.5	11.1	1538	2 T29095	cardiac muscle fac
24	399.5	11.1	1938	1 JX0178	myosin heavy chain
25	399	11.1	1931	2 A59234	slow myosin heavy
26	398	11.1	1269	2 F84730	probable myosin he
27	398	11.1	1938	2 JC5421	smooth muscle myos
28	398	11.1	1940	1 S04090	myosin heavy chain
29	398	11.1	1972	2 JC5420	smooth muscle myos

ALIGNMENTS

RESULT 1

JC5016

hyaluronan receptor - human

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 29-Sep-1999

C:Accession: JC5016

R:Wang, C.-J., Entwistle, J.; Hou, G.; Li, Q.; Turley, E.A.

(Gene 174, 299-306, 1996)

A:Title: The characterization of a human RHAMM cDNA: Conservation of the hyaluronan-b

A:Reference number: JC5016; MUID:97045829; PMID:8890751

A:Contents: breast

A:Accession: JC5016

A:Molecule type: mRNA

A:Residues: 1-725 <WAN>

A:Cross-references: GB:U29343

A:Note: it is uncertain whether Met-1 or Met-196 is the initiator

C:Comment: This receptor regulates focal adhesion turnover, and regulates human breas

C:Genetics:

A:Gene: GDB:HMWR; RHAMM

A:Cross-references: GDB:683209; OMIM:600936

A:Map position: 5q33.2-5qter

C:Superfamily: hyaluronan receptor

Query Match	100.0%;	Score 3598;	DB 1;	Length 725;
Best Local Similarity	100.0%;	Pred. No. 3e-113;		
Matches 725;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 MSFPKAPLKRFPNDPSCGAPSPGAYDVKTLEVLKGPVSFQSKQRFKQKESQNLNVDKDT 60

Db 1 MSFPKAPLKRFPNDPSCGAPSPGAYDVKTLEVLKGPVSFQSKQRFKQKESQNLNVDKDT 60

Qy 61 TLPASARKVKSESSEKSKQNDKDLKILEKEIRVLLQERGAQDRRIQDLELEKMEARL 120

Db 61 TLPASARKVKSESSEKSKQNDKDLKILEKEIRVLLQERGAQDRRIQDLELEKMEARL 120

Qy 121 NAALREKTSLSANNATLTKQLIELTRTNELLSKSFSENGNKNLRLISLLEMLKLNKREK 180

Db 121 NAALREKTSLSANNATLTKQLIELTRTNELLSKSFSENGNKNLRLISLLEMLKLNKREK 180

Qy 181 KMRGMAKQEGMEKMLQVTRSLSESQGKIAQLEGLKVSIEKEKIDSEKLEKLEKLEK 240

Db 181 KMRGMAKQEGMEKMLQVTRSLSESQGKIAQLEGLKVSIEKEKIDSEKLEKLEKLEK 240

Qy 241 ISCASDQVEKYKLDIALEENLKEKNDEILSKQSLEENIVILSKQVEDLVNKKCOLLEKE 300

Db 241 ISCASDQVEKYKLDIALEENLKEKNDEILSKQSLEENIVILSKQVEDLVNKKCOLLEKE 300

Qy 301 KEDHVRNRRENNENLNAEQNLKQKFILEQOEHEKLOQKEQLQIDSLLOQEHELSLSSHQK 360

Db 301 KEDHVRNRRENNENLNAEQNLKQKFILEQOEHEKLOQKEQLQIDSLLOQEHELSLSSHQK 360

Qy 361 LCSFQEEVVKENLFEELKQTLDELDKLOQKEQAEIRLVKQLEBEAKSRAELKLLEEK 420

myosin heavy chain  
liver stage antige  
myosin heavy chain  
restin - human  
nuclear mitotic ap  
hypothetical prote  
hypothetical prote  
hypothetical prote  
microtubule-vesicl  
myosin heavy chain  
synaptonemal compl  
embryonic muscle m  
nuclear/mitotic ap  
giantin - human  
giantin - human  
myosin heavy chain

|||||  
Db 361 LCSFOEMVKNLFEELKOTLDLQKEQAERLYKQLEBEAKSRABELKLLEK 420  
Qy 421 LAGKAELEKSSAAHTQATLLQEKYDMSVQSLDVTQAFESYKALTASEIEDLKLENS 480  
Db 421 LAGKAELEKSSAAHTQATLLQEKYDMSVQSLDVTQAFESYKALTASEIEDLKLENS 480  
Qy 481 LOEKAAGKNAEDVQHQTATLATESSENOEYVRMLDLQTKSALKETEIKETIVSFLOKITYD 540  
Db 481 LOEKAAGKNAEDVQHQTATLATESSENOEYVRMLDLQTKSALKETEIKETIVSFLOKITYD 540  
Qy 541 LQNLKQBEDPRKQLEDEBGRKAKENNTAELTEINKWRLLYBELYNKTKPFQQLDA 600  
Db 541 LQNLKQBEDPRKQLEDEBGRKAKENNTAELTEINKWRLLYBELYNKTKPFQQLDA 600  
Qy 601 FEVEKQALLNEHGAEOQLNRDSYAKLLGHONLKQKIKHVVKLKDENSQKSEVSKLR 660  
Db 601 FEVEKQALLNEHGAEOQLNRDSYAKLLGHONLKQKIKHVVKLKDENSQKSEVSKLR 660  
Qy 661 COLAKKKQSETKQBELNKVLGKIKHFDPSKAPHHESKENFALTKPLKEGNTNCRYAPMEC 720  
Db 661 COLAKKKQSETKQBELNKVLGKIKHFDPSKAPHHESKENFALTKPLKEGNTNCRYAPMEC 720  
Qy 721 QESWK 725  
|||||  
Db 721 QESWK 725

RESULT 2  
JC4298  
hyaluronan receptor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 16-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 01-Dec-2000  
C:Accession: JC4298; A42925; A41923; S21586  
R:Enright, J.; Zhang, S.; Yang, B.; Wong, C.; Li, Q.; Hall, C.L.; A, J.; Mowat, M.; G  
Gene 163, 233-238, 1995  
A:Title: Characterization of the murine gene encoding the hyaluronan receptor RHAMM.  
A:Reference number: JC4298; MUID:96011639; PMID:7590272  
A:Accession: JC4298  
A:Molecule type: mRNA  
A:Residues: 1-631 <ENT>  
A:CROSS-references: EMBL:X64550  
A:Experimental source: 3T3 fibroblast  
R:Hardwick, C.  
J. Cell Biol. 118, 753, 1992  
A:Reference number: A42925; MUID:92348516; PMID:1639856  
A:Contents: erratum  
A:Accession: A42925  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 155-376,'S',378-504,'E',506-631 <HAR>  
A:CROSS-references: GB:X64550  
A:Note: authors translated the codon AGT for residue 377 as Thr and CTA for residue 507  
R:Hardwick, C.; Hoare, K.; Owens, R.; Hohn, H.P.; Hook, M.; Moore, D.; Cripps, V.; Auste  
J. Cell Biol. 117, 1343-1350, 1992  
A:Title: Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motil  
A:Reference number: A41923; MUID:92299690; PMID:1376732  
A:Accession: A41923  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 153-376,'S',378-504,'E',506-507,'L',508-630 <HA2>  
A:CROSS-references: GB:X64550  
A:Note: this sequence has been corrected in reference A42925  
C:Comment: This protein regulates cell motility and transformation, and focal adhesion c  
C:Genetics:  
A:Gene: rhann  
A:Introns: 20/3; 54/2; 79/2; 137/3; 187/3; 259/2; 382/2; 431/2; 482/2; 515/3; 574/3; 625  
C:Superfamily: hyaluronan receptor  
C:Keywords: glycoprotein; receptor  
F:260-382/Region: 21 residue repeats  
F:516-574/Region: hyaluronan binding #status predicted  
F:575-625/Region: hyaluronan binding #status predicted  
F:116,283,304,325,346,367,398,438,619/Binding site: carbohydrate (Asn) (covalent) #statu

Query Match 54.3%; Score 1952; DB 2; Length 631;  
Best Local Similarity 64.7%; Pred. No. 1.5e-58;  
Matches 410; Conservative 51; Mismatches 85; Indels 88; Gaps 3;  
Qy 164 LRLSLEMLKRNKRETKRMGMMAKQEGMEKMLQVTQSRLESQKIAQLEGKLVYSIEKE 223  
Db 1 MRALSLEMLKRNKRETKRMGMMAKQEGMEKMLQVTQSRLESQKIAQLEGKLVYSIEKE 60  
Qy 224 KIDEKSETEKLEYEEETSCASDQVEKYKLDIAQLEENLKEKNDEILSLKOSLEENIVIL 283  
Db 61 KIDEKCEKTEKLEYIQEISCASDQVEKCKVDIAQLEEDLKEKDRILSLKOSLEENIT-F 119  
Qy 284 SKQVEDLVKCOLLEKEDEHVNRRNHNENLNAQMQLKQKFIQEQEHKILQKQELQI 343  
Db 120 SKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKLOQKELQS 179  
Qy 344 DSLQOQEKELSSSLHQKLCFSFOEMVKEKNLFEELKOTLDLDELKLOQKEQAERLVKQL 403  
Db 180 QSLQOQEKELSLRQOQLCSFQEMTSEKNVFKELKLAELDAVQOQKEQSERLVKQL 239  
Qy 404 EEEAKSRAEELKLEELKLGKAELEKSSAAHTQATLL----- 441  
Db 240 EEEKSTAEQTLRLDNLRLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 299  
Qy 442 ----- 441  
Db 300 QEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQL 359  
Qy 442 --LQEKYDMSVQSLDVTQAFESYKALTASEIEDLKLENSLQEKAAKAGNAEDVQHQI 499  
Db 360 ESVQEKYNDTAQSLRDVTAQLESYKSLTKETEDLKLENLTQEKVAMAESVEDVQOQI 419  
Qy 500 LATESSNOEYVRMLDLQTKSALKETEIKETIVSFLOKITDLQNLKQKEEDFRKQLEDE 559  
Db 420 LTAESTNOEYARMVODLQNRSTLKEEELKEITTSFLEKITDLKNLQROQDEDFRKQLESEK 479  
Qy 560 EGKRAEKENTTAELTEINKWRLLYBELYNKTKPFQQLDAFEVEKQALLNEHGAQAQBL 619  
Db 480 GKRYAEKENVMTLMEINKWRLLYDELYEKTQPOQDADAFEAKQALLNEHGATQEQBL 539  
Qy 620 NKIRDSYAKLLGHONLKQKIKHVVKLKDENSQKSEVSKLRCOLAKKKQSETKQLEELNK 679  
Db 540 NKIRDSYAQLLGHONLKQKIKHVVKLKDENSQKSEVSKLSQLVKKRQNLRLQGEJDK 599  
Qy 680 VLGIKHPDPSKAFHHESKENFALTKPLKEGNTNC 713  
Db 600 ALGIRHPDPSKAFCHASKENF---TPLKAGNPNC 630

RESULT 3  
S67593  
transport protein USO1 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein D2552; protein YDL058w  
C:Species: Saccharomyces cerevisiae  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: S67593; A38455; S30782  
R:Blöcker, H.; Brandt, P.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67587  
A:Accession: S67593  
A:Molecule type: DNA  
A:Residues: 1-1790 <BLO>  
A:CROSS-references: EMBL:Z74106; NID:g1431058; PID:e253003; MIPS:YDL058  
A:Experimental source: strain S288C  
R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.  
J. Cell Biol. 113, 245-260, 1991  
A:Title: A cytoskeleton-related gene, USO1, is required for intracellular protein tra  
A:Reference number: A38455; MUID:91185402; PMID:2010462  
A:Accession: A38455  
A:Molecule type: DNA  
A:Residues: 1-389,'TA',392-724,'S',726-1790 <NAK>  
A:CROSS-references: GB:X54378; NID:g4777; PIDN:CAA38253.1; PID:g4778







Db 611 AAQDRVLS--LETSVNELNSQLESKEKVSQD--IQIAKTELLLSAEAAKTAQADLQ 666  
Qy 413 -ELKLEKLGKAELEKSSAAHTQATLLQEKYDVMQVSLSDVTAQFESYKALTASEI 471  
Db 667 NHDOTAQNALQKHQELNKITTDQVTAQKQKHQSQLESHLKEYKYLSL-EQKT 725  
Qy 472 EDL-----KLENSLOEKAARAGNAEDVHQ-----ILATESNQ-----EYVRML 513  
Db 726 EELEGQIKLEADSLVKASKE-QALQDQOQOQRLNTDLELRATELTSQLEMEKEIVSST 784  
Qy 514 -LDLQTSALKETELKEITVTFLOKTIQDNLQNLQKQBEDFKQLEDEGRKAERKENTAE 572  
Db 785 RLDLQKKS-----FALESIKOKLT-----KOESE---KQTLKQDFETLSQET----- 823  
Qy 573 LTEEINKWRLLYEELYNKTKPFQLOLQFAVEKQALLNEHGAQOELNKIRDSY----- 626  
Db 824 -----KIOHEELNRIQTTVELQKVMEKEALMTELSTVKDKLSKVSQSLKNSKSE 875  
Qy 627 -----AKLLGHQNLKQIKKHVVKLKDNSOLKSEVSKLRQLAKKQKQSETKLOEEL 677  
Db 876 FEKENQKGAAILDEKTKELKHQLOQVOMENT-LK-EQKELKKSLEKEKEASHQKLEL 933  
Qy 678 NKV 680  
Db 934 NSM 936

RESULT 7  
S38173  
myosin-like protein MLP1 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YKR095w; protein YKR415  
C:Species: Saccharomyces cerevisiae  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 21-Jul-2000  
C:Accession: S38173; S40647; S31207  
R:Baladron, V.; Ballista, J.P.G.; Bou, G.; del Rey, F.; Esteban, P.F.; Garcia-Cantalejo, J.  
submitted to the Protein Sequence Database, March 1994  
A:Reference number: S38158  
A:Accession: S38173  
A:Molecule type: DNA  
A:Residues: 1-1875 <BAL>  
A:Cross-references: EMBL:X73541; NID:g450550; PIDN:CAA51948.1; PID:g450554  
A:Experimental source: strain S288C  
R:Bou, G.; Esteban, P.F.; Baladron, V.; Gonzalez, G.A.; Cantalejo, J.G.; Remacha, M.; J.  
Yeast 9, 1349-1354, 1993  
A:Title: The complete sequence of a 15 820 bp segment of Saccharomyces cerevisiae chromo  
A:Reference number: S40644; MUID:94205265; PMID:8154186  
A:Accession: S40647  
A:Molecule type: DNA  
A:Residues: 1-1875 <BOU>  
A:Cross-references: EMBL:X73541; NID:g450550; PIDN:CAA51948.1; PID:g450554  
A:Experimental source: strain S288C  
R:Koelling, R.; Nguyen, T.; Chen, E.Y.; Botstein, D.  
Mol. Gen. Genet. 237, 359-369, 1993  
A:Title: A new yeast gene with a myosin-like heptad repeat structure.  
A:Reference number: S31207; MUID:93247549; PMID:8483450  
A:Accession: S31207  
A:Molecule type: DNA  
A:Residues: 1-300, 'A', 302-1875 <KOE>  
A:Cross-references: EMBL:L01992; NID:g171958; PIDN:AAA34783.1; PID:g171959  
C:Genetics:  
A:Gene: SGD:MLP1  
A:Cross-references: SGD:S0001803; MIPS:YKR095w  
A:Map position: 11R

Query Match 11.6%; Score 417; DB 2; Length 1875;  
Best Local Similarity 24.1%; Pred. No. 4.9e-07;  
Matches 190; Conservative 145; Mismatches 292; Indels 162; Gaps 30;

Qy 29 LEVLKGPVSFOKSRFOQKESQNLNVND----KDTTLPASARKVKSESSEKES--QKND 82  
Db 816 LSELKETS-QKDHHIKQLEED--NNSNIEWQNKIEALKKDYESTVTSVDSKQTDIEKIQ 873  
Qy 83 KDLKILEKEI---RVLLQERCAQDRRIQD--LETELEKMEARLNAA-----LREKT 128

Db 874 YKVSLEKEIEDKIRLHTYNVMDETINDSLRLEKSKINLTADAYSQIKEYKDYETT 933  
Qy 129 SLSANNATLEKQLTELFTNELLKSKPSENGQ-KNL-----RIISLELMKLRN 176  
Db 934 SQS-----LOQTSKLDSEFKDFTNQIKNLDEKTSLEDKISLLEKQOMFLNN 981  
Qy 177 KRETKMRGMMAKQEGMEMKLOVTRSLSESGKIAQLSGKLVISLEKEXIDKS----- 229  
Db 982 ELDLQKGMKEKADFKKRISILQNNKNEEVAVKSEYKSLKIQND-LDQOTIYANTAQ 1040  
Qy 230 -----ETEKLLLEYIEEISCASQDVEKYK-----LDIAQLEENLKENDEILSLKQSL 277  
Db 1041 NNYEQELQKHADVSKTISELREQLHTYKGVQKVTNLNRDQLENALKENKSWSSOKESIL 1100  
Qy 278 ENIVILSKQVEDLVNKKCOLLEKEDHVRNRREHNENLNA-----EMQNLQAK 325  
Db 1101 EQLDLSNRIEDLSQNKLLYDQIQTAAADKEVNNSTPGFLNNILITLRERDILDTK 1160  
Qy 326 FILEQOEHEKLOQKELQIDSLLOQ-----EKELSSSLHOKLCSFOEE-----MVK 370  
Db 1161 VTVAERDAKMLRQKISLMDVELQDARTKLDNSRVEKENHSSIIQOHDDIMEKLNQLNLR 1220  
Qy 371 EKNL-FEEL-----KQTLDELDKLOQKEEQAEERLYKQLEEEAKSRAELKLEEK- 421  
Db 1221 ESNITLRNELENNNNKKELQSELQKQNVAPIESELTKALKYSMQEKEQELKLAKEVH 1280  
Qy 422 ---KGKAEAELEKSSAAHTQATLLQEKYDVMQVSL-----DVTQAFESYKALT 467  
Db 1281 RWKRSQDILEK---HQLSSSDYKLESEIENLEKENKERQGAEEKFNLRQA 1336  
Qy 468 ASETEDKLENSLSQLEKA-----AKAGNAEDVOHOILATESSNQYVVRML 513  
Db 1337 QERLTKSLQSDSLEQVNSLRDAKNVLENSANARIEELQNAKVAQGNQLEAIRKL 1396  
Qy 514 L-DLQTSALKETELKEITVTFLOKTIQDNLQNLQKQBEDFKQLEDEGRKAERKENTAE 572  
Db 1397 QEDAEKASRELQAKLEESTSVESTINGLNEEITTLKEIEKQROIQOOLQATSANEQND 1456  
Qy 573 LTEEINKWRLLYEELYNKTKPFQLOLQFAVEKQALLNEH-GAAEQELNKRIRDSVAKLLG 631  
Db 1457 LSNIVESMKKSFE--DKIK-----FIKEQTQEVNEKILEAQERLNQPSN-----IN 1501  
Qy 632 HONLKQKIKHVVKLKDNSOLKSEVS-KLR-COLAKKQKQSETKQLEELNKVLGKHFDP 689  
Db 1502 MEEKKKWE-----SEHEQEVSQKIRAEAEALKRIRLPTEEKINKIIEERKKELE 1552  
Qy 690 KAFHESKE 698  
Db 1553 KEFEKVEE 1561

RESULT 8  
T14867  
Interaptin - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000  
C:Accession: T14867  
R:Rivero, F.J.; Kuspa, A.; Brokamp, R.; Matzner, M.; Noegel, A.A.  
J. Cell Biol. 142, 735-750, 1998  
A:Title: Interaptin, an actin-binding protein of the alpha-actinin superfamily in Dic  
ts.  
A:Reference number: 218248; MUID:98365468; PMID:9700162  
A:Accession: T14867  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1738 <RIV>  
A:Cross-references: EMBL:AF057019; NID:g3549260; PID:g3549261; PIDN:AAC34582.1  
C:Genetics:  
A:Gene: abpp  
A:Introns: 173/2; 1680/1

Query Match 11.6%; Score 416; DB 2; Length 1738;

Best Local Similarity 23.5%; Pred. No. 5e-07;  
Matches 173; Conservative 157; Mismatches 278; Indels 128; Gaps 27;

QY 43 REKQKESQNLNVKDTTLTPASARKVSKSEKSKQNDKDLKILEKIRVLLQERGAQ 102  
Db 789 QYKQOOLS-SNSNIDQ-----LSTIIIEUSELKEKELNDLKEKQLOLQOEFOQL 843  
QY 103 DRRIQ-DLETELEKMARLNAALREKTSLSANNATLEKOLIELTRTNELLKSKFSENGNO 161  
Db 844 NEKNQDHQDQLELLEKQLKQLOQEVQDLNETNQSTENQNLNOLINK-----ENLNE 896  
QY 162 KNLRIILSLEMLKRLNRKRETKMGMAKQEGMEMKLOVQORSLEESQKTAQLEGLVSTIE 221  
Db 897 K-----EQELLQNLQNLQIEKIQFDQOEFKQNSINTELVNEKNEKLIQDQDYDLK 951  
QY 222 KE--KIDKSEYEKLELLEYIEELSCASQDQVEKYKLDIALEENLKEKNDLILSKOSLEEN 279  
Db 952 QNRSNDKEND-LIE-----KENOLKSTONELNOLIEKNEKSHKBEQQLKQOSIEND 1003  
QY 280 IVILSKQVEDLNVKQLEK-----EKEDHYNRNRHNEHNLNAEMNLKQKFILEQ 330  
Db 1004 LIEKENQIQL--QSOLNROQOQSNOLSEKQDQNLQLEKQ--FDQEQQLKQOSIEND 1060  
QY 331 --OEHEKIQOQLDLSLQKELSSSHQKCLSFQSEEMVKEKNLFEELKQTLDELK 388  
Db 1061 LFEKENQIQOQLSQNLQNEQOQSNOLSEKQDQNLQLEK--NESDQKEQQLKQOSIEND- 1117  
QY 389 LQKQEQARLVKQLEEEAKRAE-----ELKLEELKKGKEAELEKSSAAHTQATLL 442  
Db 1118 LIEKENQIQOQLQNEQQLQSEIVSINDKILEKQKQOSDQLLNDEKQOQDKOL 1177  
QY 443 QEK---YDSM-----VQSLVDVTAQFESYKALTAIEIEDLKLENSSLQEK 484  
Db 1178 QDKQIEFDQQLTFQKFNKDKDSQFIQDDQKQLOSTQ-----QDLNOLQKQOENKEKQ 1233  
QY 485 AKAGKNAEDVOHILATESNQEVYVRLMLDLQTSALKETKEITVTSFLOKITDLO-- 542  
Db 1234 LSEKDEKLSIQFQNEQEKQLSEKDEKLSQSQNLQNLQNLQNDQNEKQVQFSEKDEKLSI 1293  
QY 543 ---NOLKQDEDFRQKDEDEGRKAEKENTTAETTERINKRLWLYEELYNTKPFQLOL 598  
Db 1294 QDLNOLQKQOENKEKQ-----SEKDEKLSQSQDQNLQ-----LNDQIKNEKQL--- 1340  
QY 599 DAFEVEKQALLNEHGAQBLNKRIRDSYAKLLGHQNLKQIKHYVVKLDKENSQKSE--V 656  
Db 1341 ---EKEEQLLKIQDQNDQSQQLQLEKLESEKENQQLQKQENIQNLQOQOSNEII 1397  
QY 657 SKLQOLAKKKQSE-----TKLQBEL-----NKVLGI-----KHFD 687  
Db 1398 QQLKQDLKQKQOQOQOQENNEKIEIRLQIEIRQLKQOQOQIDQSELSNKEIKIQTQKQEF 1457  
QY 688 PSKAFHESKENFALK 703  
Db 1458 ---QLSHNRSKOQLHLQ 1471

RESULT 9  
A59252  
myosin heavy chain, nonmuscle, form IIB - human  
N:Alternate names: myosin type 10; NMHC-B  
N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Homo sapiens (man)  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 19-Apr-2002  
C:Accession: A59252; B61231; G02055  
R:Phillips, C.L.; Yamakawa, K.; Adelstein, R.S.  
J. Muscle Res. Cell. Motil. 16, 379-389, 1995  
A:Title: Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and analysis  
A:Reference number: A59252; MUID:96025307; PMID:7499478  
A:Accession: A59252  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1976 <SIM>  
A:Cross-references: GB:M69181; NID:g641957; PIDN:AAA99177.1; PID:g641958

A:Experimental source: clone lib Lambda Zap II adult human T-cell library; cell line  
A:Note: between nucleotides 1942-1943 in mRNA encoding human brain MHC-B there is an  
R:Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Gdula, D.; Adelstein  
Circ. Res. 69, 530-539, 1991  
A:Title: Human nonmuscle myosin heavy chains are encoded by two genes located on diff  
A:Reference number: A61231; MUID:91316803; PMID:1860190  
A:Accession: B61231  
A:Molecule type: mRNA  
A:Residues: 63-237, 'K', 239-664, 'L', 666-722 <SI2>  
A:Cross-references: GB:M69181; NID:g641957  
R:Weir, L.  
submitted to the EMBL Data Library, August 1995  
A:Reference number: H00753  
A:Accession: G02055  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-81 <WEI>  
A:Cross-references: EMBL:U34304; NID:g1143217; PIDN:AAA84880.1; PID:g1143218  
C:Genetics:  
A:Gene: GDB:MYH10  
A:Cross-references: GDB:127350; GDB:G00-127-350; OMIM:160776  
A:Map position: 17p13-17p13  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleo  
F:88-771/Domain: myosin motor domain homology <MMO>  
F:178-185/Region: nucleotide-binding motif A (P-loop)  
F:559-572/Region: actin binding #status predicted  
F:633-647/Region: actin binding #status predicted  
F:123/Modified site: N6,N6-trimethyllysine (Lys) #status predicted  
F:184/Binding site: ATP (Lys) #status predicted  
F:701,711/Active site: Cys #status predicted

Query Match 11.5%; Score 414; DB 2; Length 1976;  
Best Local Similarity 23.4%; Pred. No. 6.5e-07;  
Matches 199; Conservative 150; Mismatches 310; Indels 190; Gaps 29;

QY 32 LKGPVSFQSRQKQKESQNLNVKDTTL-PASARKV-----KSS 72  
Db 1074 LQAQIDELKQLAKKEELQALARGDDDETLHKNNALVVRLOAQIAELQEDFESEKAS 1133  
QY 73 ESKRESQND--KDKLILEKEIRVLLQERGAQ-----DDR 105  
Db 1134 RNKAEKQKRDSEELAEKLELTDLTAAQOELTRKEQEVAEKKALEBETKNHEAQ 1193  
QY 106 IQDLE---TELEKEARLNAALREKTSLSANNATLEKQTLTRTNELLKSKFSENGNO 161  
Db 1194 IQDMQRHATALEELSEQLQAKKFKANLEKNKQGLTDNKLACQVAVLQVKAESHK 1253  
QY 162 K-----NLRI-LSLEMLKRLNK-----RETKRMGMAKQ--GM 192  
Db 1254 RKKLDAQVOELHAKVSEGRRLRVLAELAKSKLQNLNDNVSTLLEAEKKGKIFAKDAASL 1313  
QY 193 EMKLQVQTSRLEESQKTAQLEGLVSTIEKEKIDSEKTEKLELLEYIEELSCASQDQVEYK 252  
Db 1314 ESQLODTQLEETROKLNLSRRIRQLEEK-----NSLQEQEKEEARKNLEKQV 1366  
QY 253 LDIAQLEENLEKNDLILSKQSLPE-----NIVILSKOVEDLNVKCOLLEKEKEDHV 305  
Db 1367 LAIQSLQADTKKKVDDDLGTTIESLEAKKLLKDLAEALSQRLEEKALAYDKLETK- 1422  
QY 306 NRNRHNNLNAEMNLKQKFILEQOHEKLEKQKELQIDLSLQOKEKELSSSHQK----- 360  
Db 1423 NRIQQLDQDLVLDLH-----QRQVASNLEKKQKFKDQLLAEBKSSISARYAEERDRAE 1475  
QY 361 -----LCSFQEMVKEKNLFEELKQTLDELKQOKEQEARLVKQLEBEAK 408  
Db 1476 AEAREKETKALSARALEAELEAKPEFRQNKQLRADMEDLMSSKDDYGVKNVHELEKSR 1535  
QY 409 SRAEELKLEELKKGKEAELEKSSAAHTQATLLQLEKVDYMSVQSLDVTQAFESYKALTA 468  
Db 1536 ALQEQVEEMTQLEBELEDELOATEDAKLRLVNVQAKQAQFERDQTRDQNEEKRLLI 1595  
QY 469 SEIEDL--KLENSSLQEKAAKAGK-----NAEDVOHQILATESNQEVYVRLMLDLQTKSA 521

Db 1596 KQVRELEAELEDERKQRALAVASKKMEIDLKLEAQTEAANKARDEVIKQRLKLAQMK 1655  
Qy 522 LKETEIKETIYS---FLQ-----KTTDLQNLKQOE-----EDFRKQLED 558  
Db 1656 DYQRELEARASRDEIFAQSVESSEKKULSLAEITLQOEELASSERARRHAEQERDELAD 1715  
Qy 559 E-----EGRKA---EK-----ENTTAELTEEINKWRLLYEELYNKTKPFQQLDFAFEVEKQA 607  
Db 1716 EITNSASCKSALLDEKRLEARIAQLEEEEOSNMELLNDRFRKTTIQLQVDTLNAE--- 1772  
Qy 608 LLNEHGAQOEQLKIRDSYAKLGHON--LKQIKHV-----VKLKDNSOLSKSEVSKLR 660  
Db 1773 LAAERSAAQKS-----DNARQOLERONKELKAKLQELGAVKSKFKATISALEAKIGOLE 1827  
Qy 661 COL---AKKKOSETKLOEELKNVL-----GKHPDPSKAFHESKENF-ALKTPLK 707  
Db 1828 EQLQOEAKERAANKLVRRTEKKLKEIFPMQVEDERRHADQYKEQMEKANARKOLKROLE 1887  
Qy 708 EGNNTCYRA 716  
Db 1888 EAEBEATRA 1896

RESULT 10  
A26655  
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)  
N:Contains: myosin Atpase (EC 3.6.4.1)  
C:Species: Dictyostelium discoidaeum  
C>Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 19-Apr-2002  
C:Accession: A26655; A24728; S00250  
R:Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.  
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986  
A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum  
A:Reference number: A26655; MUID:87092266; PMID:3540939  
A:Accession: A26655  
A:Molecule type: DNA  
A:Residues: 1-2116 <WAR>  
A:CROSS-references: GB:M14628; GB:M11938; NID:g167834; PIDN:AAA33227.1; PID:g167835  
R:DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985  
A:Reference number: A24728; MUID:86016788; PMID:3901008  
A:Accession: A24728  
A:Molecule type: mRNA  
A:Residues: 2035-2116 <DEL>  
R:Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.  
FEBS Lett. 227, 71-75, 1988  
A:Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium  
A:Reference number: S00250; MUID:88112226; PMID:2828113  
A:Accession: S00250  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1734-1893 <WAG>  
C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphatase  
F:1-818/Domain: globular head <HEAD>  
F:89-747/Domain: myosin motor domain homology <MMOT>  
F:179-186/Region: nucleotide-binding motif A (P-loop)  
F:819-2116/Domain: alpha-helical rod <ROD>

Query Match 11.5%; Score 414; DB 2; Length 2116;  
Best Local Similarity 22.3%; Pred. No. 7e-07;  
Matches 191; Conservative 178; Mismatches 275; Indels 214; Gaps 36;  
Qy 8 LKRFNDPSGCAID-VKTLVLPKVPVFSQK-----SQRFOOKESKONLVNDK--- 58  
Db 959 MKRVND-----GQSDTISRLEIKID--ELQKEVEELTESFSBESKDKGRVLEKTRVRL 1008  
Qy 59 -----DTTLPASARKVSKESKESKNDKDLKILEKEIRVILQERGAQDRRIQDLETLEL 113  
Db 1009 QSELDLTVRLDSETKOKSELLRQKKLEELKQVQALAAETAAKLAQEAANKLQGEY 1068

Qy 114 EKMEARLNAALREKRTSLSNATTLEKQILTRTNELLKSKFSENGKNQNLRIISLELMK 173  
Db 1069 TELNEKFNSEVTARSNEVSKSKTTLESQLVAV--NNELDEEKNRDALKKKKALDAMLEE 1126  
Qy 174 LRNRETKMRGMAKQEGEMK-----LQVTQBS--LEESQGIQAEGLGLVSIERKIDEX 228  
Db 1127 MKDLES-----TGGEKSKLYDLKVKQFESDMELRNQISELQSTIAKLEKTKSTLE 1177  
Qy 229 SETKLLYIEEISCASQVQEKY----LDI-----AQLEENLKEKNDEILSKQSLLENI 280  
Db 1178 GEVARLQELAEALAKSNVEKQKKVVDLEDKSAQLAETAAK-QALDKLKKLEBEL 1236  
Qy 281 VILSKQVEDLVKCOLLEKEKEDHYNNRHNENINAEOMNLKQKFIIEQQEHEKLOQKE 340  
Db 1237 SEVQTLSEANK-----NVNSD--STNKHLETFSNNLKLEAEAQAKAQALEKKR 1285  
Qy 341 LQIDS-----LLOQEKELSSLLHOKLCSQFQEMVKNLFEELKQTLDELKLOQKE 393  
Db 1286 LGLESELKHVNEQLEEEKQKESNEKRVKVDLEKEVSELKDQIEEVASKKAVTEAKNKE 1345  
Qy 394 BOAERLVKOLEEAKSR---AEELKLLEK-----LKGKEAELEKSSAAHTQATLLQ 443  
Db 1346 SELDEIKROYADVSSRDKSVQQLTKLQAKNEELNNTAEAEQGLDRAERSKKAEFDLE 1405  
Qy 444 EKYDSMVQSLDVTQ-OFESYKAL-----TASEIEDLK----- 475  
Db 1406 E---AVKNLEETAKKVKAEKAMKKAETDYRSTKSELDDAKNVSSSEQVVOIKRLNEELS 1461  
Qy 476 -----LE-----NSSLOEK-----AAKAGKNAEDVQHOILATES 504  
Db 1462 ELRSVLEADRCNSAIKAKKTAESALESLEKDEIDAANNKAKAEKRSKELEVRVALEEE 1521  
Qy 505 S-----NOEYV-----RMLDLOTKSALKTEIKE----- 529  
Db 1522 SLEPKSGTVNVFEIRKKDAEIDDLRAERDRETESIKSDEDKKNTWKQFADLEAKVEAQ 1581  
Qy 530 ---ITVSFLQK-----ITDLQNLKQOE-----EDFRKQLEDE--EGRKAKENTTABEL 574  
Db 1582 REVVTIDRLKKLESIDIDLSTQDLTETKSRIKIEKSKKLEQTLAERRAAEGSSKAAD 1641  
Qy 575 EEINKWRLLYEEELYNKTKPFQQLDFAFEVEKQALLNEHGAQOEQLNKIRDSYAKLGHON 634  
Db 1642 EEIRK-----QVMQEVDELRAQLDS-----ERAAALN-----ASEK--KIKSLVAEV---DE 1682  
Qy 635 LKQIKHVVKLKDE----NSOLKSEVSKRLQAKQKQSETKLOEELNKLVLGKIKHFDPSK 690  
Db 1683 VKEQLEDEILAKDLKLVKAKRALEVELEEVROQLEEEEDSRSEL-EDSKRRLTTEVEDIKK 1741  
Qy 691 AFHESKENFALKTPKE 708  
Db 1742 KYDAEVEQNTKLDEAKKK 1759  
RESULT 11  
A33977  
myosin heavy chain, nonmuscle - chicken  
N:Contains: myosin Atpase (EC 3.6.4.1)  
C:Species: Gallus gallus (Chicken)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 19-Apr-2002  
C:Accession: A33977; S06116; A34422  
R:Shohet, R.V.; Conti, M.A.; Kawamoto, S.; Preston, Y.A.; Brill, D.A.; Adelstein, R.S.  
Proc. Natl. Acad. Sci. U.S.A. 86, 7726-7730, 1989  
A:Title: Cloning of the cDNA encoding the myosin heavy chain of a vertebrate cellular  
A:Reference number: A33977; MUID:90046668; PMID:2813355  
A:Accession: A33977  
A:Molecule type: mRNA  
A:Residues: 1-1959 <SHO>  
A:CROSS-references: GB:M26510; NID:g212382; PIDN:AAA48974.1; PID:g212383  
R:Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Masaki, T.  
Eur. J. Biochem. 184, 611-616, 1989  
A:Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed  
s.  
A:Reference number: S06116; MUID:90032648; PMID:2806244

A:Accession: S06116  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 716-1008 <KAT>  
A:Cross-references: GB:X17589  
A>Note: this translation is not annotated in GenBank entry GGMHCFMHA, release 114  
R:Hodge, T.P.; Cross, R.; Kendrick-Jones, J.  
J. Cell Biol. 118, 1085-1095, 1992  
A:Title: Role of the COOH-terminal nonhelical tailpiece in the assembly of a vertebrate  
A:Reference number: A43422; MUID:92381096; PMID:15112291  
A:Accession: A43422  
A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1900-1959 <HOD>  
A:Experimental source: brush border  
A>Note: sequence extracted from NCBI backbone (NCBIP:111947)  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide  
F:84-764/Domain: myosin motor domain homology <MMOT>  
F:174-181/Region: nucleotide-binding motif A (P-loop)  
F:552-565/Region: actin binding #status predicted  
F:626-640/Region: actin binding #status predicted  
F:837-1936/Domain: coiled coil #status predicted <COI>  
F:837-1277/Region: S2  
F:1278-1959/Region: light meromyosin  
F:1937-1959/Domain: carboxyl-terminal <CBT>  
F:125/Modified site: N6,N6-trimethyllysine (Lys) #status predicted  
F:180/Binding site: ATP (Lys) #status predicted  
F:694,704/Active site: Cys #status predicted

Query Match 11.4%; Score 409; DB 1; Length 1959;  
Best Local Similarity 23.0%; Pred. No. 9.5e-07;  
Matches 189; Conservative 159; Mismatches 289; Indels 184; Gaps 32;

Qy 37 SFQSKORFQKQESKQNLNVKDTTLPSARVKVSESKESQK--NKDKLKILEKEIRV 94  
Db 873 TFO-AQLMAEKQLOLQAEALCAEAETRLAKKQLEETCHDLAEVVEEERC 931  
Qy 95 --LLEAGQADRRIDOLETELEKME-ARLNAAAREKTS-----LSANNATLEKOLIET 145  
Db 932 QHLQAEKKKMQNIQLEOLEEESARQKQLEKVTYFAKLLKLEEDVIVLEDONLKLA 991  
Qy 146 RTNELLSKFSF-NGNQNLRLSLLEMLKLRNRE---TKRMGMMAKQMGEMKLVQTVR 201  
Db 992 KEKKLLDRMSEFTNLATEEESKSLAKKNKHEAMITDLERLRREKQKOLEKTRR 1051  
Qy 202 SLE-----ESQKIA-----QLEGLVSTKE-----KIDE-KS 229  
Db 1052 KLEGGSSDLHDQIAELQIAELKQTSKKEEELQALAEVVEEAAQKNMALKKIRELS 1111  
Qy 230 EPEKLEYTEITSASDOVEKYKLDIAOLEENLEKNDIEL-----SLKQSLLENTVI 282  
Db 1112 QITEQLQEDLESASRNKAQKQKRDGLGEELEALTELEDLSTAAQELSKRQEVTV 1171  
Qy 283 LSKQVED-----LNVKCOLLEKEKEDHVN 306  
Db 1172 LKKTLEDAKTHAQIQEMRKHSQAIBELAQLEQTKRVKRVANLEKAKQALESERAEISN 1231  
Qy 307 R-----NREH-NENLNAEMNLKQKFI-----LEQEHKLOQKELQIDS---LLQQ 349  
Db 1232 EYKVLQKGDAEHRKKRVDAQLQELQKFTGERVKVTKLEAEARNVKQLQVLDNVTVGLNQ 1291  
Qy 350 EKELS-----SSLRQLKCSFQEMVKEKNL---FEPELQTLDELKLOQKEQAERL 399  
Db 1292 SDSKSIKLAKDFSALQSLODQQLQELQETRLKLSFSLKQTEDEKNALEKOLEEEREA 1351  
Qy 400 VKQLEE-----EAKSRAE-----LKLEELKKEAELEKSSAAHTQ---ATLLL 442  
Db 1352 KRNLEKQISVLQQAQVAREARKMDGGLGCLTAEEAAKKLQKLDLESLTRYBEKTAAYDKL 1411  
Qy 443 QEKYDSVMQSLDVTVAQFESYKALTASTIE-----DLKLENSLOEKAA---KAGKN 491  
Db 1412 EKTTRLQOELDDIAVDLD-HQRTQVSNLEKKQKKQKFDQGLLAEEKNISAKYAEERDRAE 1470

Qy 492 AEDVQHQLIATESNQEVYRMLLDLQTKSALKETBIKITVS-----FLQKITDLQNL 545  
Db 1471 AREKETKALSARALEAELEQKAELEVRNQKQPTMEDLMSSKDDVGKSVHELEKRAL 1530  
Qy 546 KOQEDFRKQLEDEGRKAENKTAETEEINKWR-----LLYEELYNKTPFQQL- 596  
Db 1531 EQQVEEMKTOLEDELEQATED--AKLRLEVNOQAMKAQFQDRLGRDEQNEERKQLI 1588  
Qy 597 -OLDAFEVEKQALLNEHGAQPOLNKRDSYAKLGH-----QNLKQKIKHVVKLKDENS 650  
Db 1589 ROVREMEVELEDERKQSTAVAAKKLELDLKDLESHIDTANKNDEAIKHVRKLOAQHK 1648  
Qy 651 QLKSEVSKLRQ----LAKKKQSETK-----LQEEEL 677  
Db 1649 DYMRELEDTRTSREILAQAKENKKLKSMEAEMLQLOEEL 1689

RESULT 12  
B43402  
myosin heavy chain-B, neuronal - chicken  
N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Gallus gallus (Chicken)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 19-Apr-2002  
C:Accession: B43402; A43402  
R:Takahashi, M.; Kawamoto, S.; Adelstein, R.S.  
J. Biol. Chem. 267, 17864-17871, 1992  
A:Title: Evidence for inserted sequences in the head region of nonmuscle myosin spec1  
yosin.  
A:Reference number: A43402; MUID:92388144; PMID:1355479  
A:Accession: B43402  
A:Molecule type: mRNA  
A:Residues: 1-2007 <TAK>  
A:Cross-references: GB:M93676; NID:g212448; PIDN:AAA48988.1; PID:g212452  
A>Note: the sequence of residues 212-221 and 632-652 and the corresponding nucleotide  
A:Accession: A43402  
A:Molecule type: mRNA  
A:Residues: 1-211,222-631,653-2007 <TA2>  
A:Cross-references: GB:M93676; NID:g212448; PIDN:AAA48985.1; PID:g212449  
A>Note: sequence extracted from NCBI backbone (NCBIN:112864)  
C:Comment: Alternatively spliced segments 1 and 2 are found exclusively in nonmuscle  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methyl  
F:1-2007/Product: myosin heavy chain-B, neuronal #status predicted <MNS>  
F:1-211,222-631,653-2007/Product: myosin heavy chain-B, nonmuscle #status predicted <  
F:88-802/Domain: myosin motor domain homology <MMOT>  
F:178-185/Region: nucleotide-binding motif A (P-loop)  
F:212-221/Region: alternatively spliced segment 1 #status experimental  
F:559-593/Region: actin binding #status predicted  
F:632-652/Region: alternatively spliced segment 2 #status experimental  
F:692-714/Region: actin binding #status predicted  
F:875-2007/Domain: coiled coil #status predicted  
F:875-1315/Region: S2  
F:1316-2007/Region: light meromyosin  
F:129/Modified site: N6,N6-trimethyllysine (Lys) #status predicted  
F:184/Binding site: ATP (Lys) #status predicted  
F:732,742/Active site: Cys #status predicted  
F:1954/Binding site: phosphate (Thr) (covalent) #status predicted  
F:1987/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 11.4%; Score 409; DB 1; Length 2007;  
Best Local Similarity 23.6%; Pred. No. 9.7e-07;  
Matches 199; Conservative 152; Mismatches 314; Indels 178; Gaps 30;

Qy 32 LKGPVSFOKSORFQKQESKQNL-NVKDVTTLPSARKV-----KSS 72  
Db 1105 LQAQIEELKTLQAKKEELQALAEVQKNNALKVIRELQALQIAELEDSEKAS 1164  
Qy 73 ESKESQKND--KDLKILEKIRVLLQERGAQ-----DRR 105  
Db 1165 RNKAQKQKRDLSLEALKTELEDLTDTAAQQLRTKREQVEALKAIBETKNEAQ 1224  
Qy 106 IQDLE----TELEKWEARLNAAAREKTSLSANNATLEKQLIELTRTNELLSKFSENGQ 161

Db 1225 IQEIRQRHATALELSQLEQAKRFKANLEKNQGLSDNKLACVAVKVLQOVKAESHK 1284  
Qy 162 K-----NLRI--LSLEMLKRNK-----RETKMRGMMAQOE--GM 192  
Db 1285 RKKLDAQVQELTAKVTGERLVELAELAKANQLONELDNVSSLLAEABKKGIKFAKDAASL 1344  
Qy 193 EMKLQVTRSORLESQGIQALQEGKLVIKIEKIDSEKTEKLLVEIEICASDOVEKYK 252  
Db 1345 ESQLODTQELLOETRQKLNLSRIROLEEB-----NNLQEQEVEEAEARNLEKQM 1397  
Qy 253 LDI--AQLEENLKERNDILSKQLEENIVILSKQVEDLVNKKQLLEKEKEKEDHVNRRNH 311  
Db 1398 LALQAALAE-AKKVDDDLGTIEGLEENKKLLKDMESLS---QRLB-ERAMAYDKLEKT 1452  
Qy 312 NENLNAEMONLQKQFILEQOEHEKLOQKELQIDSLLOQEBELSSLLHOK----- 360  
Db 1453 KNRLOQELDDLVMDLHORQIVSNLEKQKQKQFDOQMLAEEKNISARYAEERDRAPAEAREK 1512  
Qy 361 -----LCSFOEEMVKEKNLFEELKQTLDELKQOEQAERLVKQLEEEAKSRAEEL 414  
Db 1513 ETKALSARALEALEAKEFERQNKQLRADMEDLMSKDDVGKNVHIELEKSKRTLEQOV 1572  
Qy 415 KLEEKILGKBAELEKSSAAHTQATLLQEKYDSMWQSLVEDVTQAQFYSYKALTASEIDL 474  
Db 1573 EEMRTQLELEDELOATEDAKRLLEVNNQAMKAQFERDLQARDPQNEEKKRMLVKQVREL 1632  
Qy 475 --KLENSLSQEAKAQK-----NAEDVQHOILATESNQYEVVRLDLQTKSALKETEI 527  
Db 1633 EAELEDERKORALAAKMMQMDKQLEGIEAANKARDEAIKQRLKLAQMKDYQREL 1692  
Qy 528 KEITVS---FLQ-----KITDQNLQKQOEEDF-----RKQLEDE----- 559  
Db 1693 EEARSRDEIFAQSEKSKLGLAEILQLOEFASFASERARRHAEQERDELADEIANS 1752  
Qy 560 BGRKA---EK---ENTTAETEEKNRWLLYELYNKTKPPQLQDLDAFEVEKQALLNEHG 613  
Db 1753 SGKSALLDEKRLREARIAQLEEELEEQSNMELNFRKTTLQVDTLNSB---LAGERS 1809  
Qy 614 AAQEQNL-----KIRDSYAKLGHQNLKQIKHVHKLKDNSQKSEVSKRLCOL--- 663  
Db 1810 AAQSEANARQOLRQNKELAKL---QELGSEVKS---KFKATISTLEAKIAQLEEQLEQE 1864  
Qy 664 AKKQSETKLOEELNKVL-----GIKHFDPSKAFHHESKNF-ALKTPLEKGTNC 713  
Db 1865 AKERAANKLVRRTTEKLEKVEFMQVEDERRHADQYKEOMEKANARMQKLRQLEAEAEA 1924  
Qy 714 YRA 716  
Db 1925 TRA 1927

RESULT 13

A36014  
myosin heavy chain, nonmuscle - fruit fly (Drosophila melanogaster)  
N:Contains: myosin APPase (EC 3.6.4.1)  
C:Species: Drosophila melanogaster  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 19-Apr-2002  
C:Accession: A36014; B36014  
R:Ketchum, A.S.; Stewart, C.T.; Stewart, M.; Kiehart, D.P.  
Proc. Natl. Acad. Sci. U.S.A. 87, 6316-6320, 1990  
A:Title: Complete sequence of the Drosophila nonmuscle myosin heavy-chain transcript: cd  
A:Reference number: A36014; MUID:90349606; PMID:2117279  
A:Accession: A36014  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-2017 <KET>  
A:Cross-references: GB:M3012  
C:Genetics:  
A:Gene: FlyBase:zip  
A:Cross-references: FlyBase:FBgn0005634  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated

F;1-2017/Product: myosin heavy chain, form I #status predicted <MAI>  
F;46-2017/Product: myosin heavy chain, form II #status predicted <MA2>  
F;135-815/Domain: myosin motor domain homology <MMOT>  
F;225-232/Region: nucleotide-binding motif A (P-loop)  
F;598-631/Region: actin binding #status predicted  
F;705-727/Region: actin binding #status predicted  
F;888-2017/Domain: coiled coil #status predicted <COI>  
F;888-1328/Region: S2  
F;1329-2017/Region: light meromyosin  
F;176/Modified site: N6,N6-trimethyllysine (Lys) #status predicted  
F;231/Binding site: ATP (Lys) #status predicted  
F;745,755/Active site: Cys #status predicted

Query Match 11.3% Score 408; DB 1; Length 2017;  
Best Local Similarity 24.7% Pred. No. 1.1e-06;  
Matches 209; Conservative 134; Mismatches 286; Indels 218; Gaps 35;

Qy 29 LEVLGPVSP-QKSQRFKQKESQNLN-----VDKDT-----TLPA 64  
Db 889 LEVTKQEEKLVQKEDELKQVREKLDTLAKNTQVEYKQVALVEKTTLAEQLQAEIELCA 948  
Qy 65 SARKVKS-----SESKKSQKNDKDKILEKEIRVLL--OERGAQDRRIQDLETELEKMEA 118  
Db 949 EAEESRRLMARKOELEDMQOELETRIEEERERVIALGGEKKLELNLQIDLEQLEEEA 1008  
Qy 119 RLNAALREKTSLSA-----NNATL-EKOLIELTRTNELLSKFSSENGNQK 162  
Db 1009 ARQKQLEKQVOLDKIKKYBEDLALTDQONOKLLKEKLLB-ERANDLSQTLAEEEKAK 1067  
Qy 163 NLRILSLEMLKRNKRE---TKMRGMMAQEGMEMKLVQVTRSL-----ESQ 208  
Db 1068 H-----LAKLAKHEATITELERLHKDQOQESDRSKRTIETEVADLKEQLNERRV 1120  
Qy 209 KIAQLEKLVSIKES-----KIDEKSETE-----KLEYTEEISCASDQ 247  
Db 1121 QVDEMQALAKREBELTQTLRIDDESATKATAQAKQARELESQALAEQEDLEAKAARAK 1180  
Qy 248 VEKVKLDIAOLEENLK-----EKNDIEILSKOSLENTV----- 281  
Db 1181 AEKYRDLSEELAEKLNELSLDTTAAQQLRSKRQELATLAKSLEETVNHGVLAD 1240  
Qy 282 -----ILSKQVEDLVNKKQLEKEKEDE-----HVNRRNHNENLN-- 316  
Db 1241 MRKHSQELNSINDQENLNRKAKTVLEKAKCTLEAENADLATELSRVNSSRQENRRRQ 1300  
Qy 317 -----AEMQNLKQFILEQOEHEKLOQKELQIDSLLOQEKELSSLLHOKLCSF 364  
Db 1301 AESQIAELQVLAETIERARSEL---OEKCTKLOQEAENITNOL--EAEELKASAAVKSASN 1356  
Qy 365 QEEMVKE-KNLFEELKQTLDELKQOEQAERLVKQLEEE--EAKSRABELKLE--- 418  
Db 1357 MESQLETAQQLLEETQKQLGLSSKLRQISEKALQOELEDEDEAK-RNYERKLAEVTT 1415  
Qy 419 --EKLGKAELEKSSAAHTQATLLQEKYDSMWQSLVEDVTQAQ--FESYKALTASEID 473  
Db 1416 QMOETIKKAEEDADLAKLEBGGKRLNKDIEALERQVVKELIAQNDRLDKSKKIQSELED 1475  
Qy 474 KLENSLSQEAKAAGNAEDVQHOILATESNQYEVVRLDLQTKSALKETEKEITVS 533  
Db 1476 ATIELEAQRQTKVLEKQKQNF-KILAEKAKISEQIAQERDQTAAREA-REKETKVLVS 1533  
Qy 534 -----FLQKITDQNLQKQOEDEFRKQLEDEEGRKAEKENTTAELTEINKRWLLYEBLY 588  
Db 1534 RELDEAFDKTEDLENK-----RKTQNLDELDTANTQGTADKNVHELEKAKRALESOL 1585  
Qy 589 NKTK-----PFQQLQIDA---FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQ 637  
Db 1586 AELKAQNEELEDLQLTEDAKLRLEVNNQALRSQFE-----ROLLAK---EEGAE 1633  
Qy 638 KIKHVYK-LKDENSQKSEVSKRLQCLAKKQSETKLOEELNKVGLGKHFDPSKAFHHES 696  
Db 1634 KRRGLVQRLDLEDEERKQRTAAVASKKKEGLDKE-----IETTMEHMKV 1683



QY	697	KENFALK	703
		:	
Dd	1684	KED-ALK	1689

## RESULT 14

S61477 myosin II heavy chain, non-muscle - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C:Date: 19-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 02-Feb-2001  
C:Accession: S61477; S65349  
R:Mansfield, S.G.; Al-Shirawi, D.Y.; Ketchum, A.S.; Newbern, E.C.; Kiehart, D.P.  
J. Mol. Biol. 255, 98-109, 1996  
A:Title: Molecular organization and alternative splicing in zipper, the gene that encodes  
A:Reference number: S61477; MUID:96144835; PMID:8568878  
A:Accession: S61477  
A:Molecule type: DNA  
A:Residues: 1-2057 <MAN>  
A:Cross-references: EMBL:U35816  
R:Mansfield, S.G.; Al-Shirawi, D.Y.; Ketchum, A.S.; Newbern, E.C.; Kiehart, D.P.  
Submitted to the EMBL Data Library, September 1995  
A:Reference number: S65349  
A:Accession: S65349  
A:Molecule type: DNA  
A:Residues: 1-1908, 'NL', 1911-2057 <MAW>  
A:Cross-references: EMBL:U35816; NID:g1141789; PIDN:AAB09049.1; PID:g1572481  
C:Genetics:  
A:Gene: zip  
A:Cross-references: FlyBase:FBgn0005634  
A:Introns: 10/3; 90/3; 162/3; 215/1; 255/3; 263/3; 303/3; 334/3; 1248/1; 1349/3; 1536/3;  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: alternative splicing; ATP; nucleotide binding; P-loop  
F:135-855/Domain: myosin motor domain homology <NMOT>  
F:225-232/Region: nucleotide-binding motif A (P-loop)



Qy 289 ----DLNVKCOLLEKEKEDHVR-----NREH-NENLNADMQLKOKFI-----LEQ 330  
Db 1210 RLKGNLEKAKALEGSENELANVKTLLQKGDSEHKKRKKVEAQLOELQVKVTEGDRVS 1269  
Qy 331 QEHEKLAQKELQID---SLQOEKELSSHQKLCFQOEMVKEKNLFEELKQTLDELQ 387  
Db 1270 ELSEKANRLQVELDNVNSLLSQSDSKSKILGKDFSTLESQFQDAQELQEBETROKLSFST 1329  
Qy 388 KLOQKEQAERLVKOLEEAKSR-----AEEL-KL 416  
Db 1330 KLQMEDEKNGLLQLEEEBAKKNLCKQISTLOSQMTMKKMDENVGSLVTEELKK 1389  
Qy 417 LEEKLGKEAELEKSSAAHTQ----ATLLQEKYD-----SMVQSLEDVTAQFESY 463  
Db 1390 LQKLEAVNQFFEKAAAYDKLEKTKRLOQELDDISVDLDHQRIQIVSNLEKKQKKFDQL 1449  
Qy 464 ----KALTA-----SEIEDL 474  
Db 1450 LGEKKAISAKYADRRDRSEAEAREKETKALALARALEEALEAKAELERLNKQLTEMEDL 1509  
Qy 475 -----KLENS--SLOEKAAGNAEDVQHQILATESNQYVYRMLDLQTKS 520  
Db 1510 VSSKDDVCKSVHELEKSKRALEQQAEMKTOLEEELEQATEDAK---LRLEYNLQAMK 1566  
Qy 521 ALKETEIKEITVSLQKITDILQNLKQOE---EDFRKQ-----LEDEEGR- 562  
Db 1567 AQFERDLQGRDEQSEDKKQLVRQKEMEALEDERKQRSLAVAARKKLEMDLKDLEGQI 1626  
Qy 563 ---KAEKENTAELE-----EINKWRLLYBELYNKTKPFQOLDAFEVEKQALL 609  
Db 1627 DSSNKNREDAIKQLKLAQIKDYQRELEDTRASRDDILSQSKESKKLKSMEAEMIHMQ 1686  
Qy 610 NEHCA-----AQEOLNKIRDSYAKLLGHQNLKQKIKHVVKLKDENSOLKSEVSKLR-- 660  
Db 1687 EELAAAEGRKQAQOQERDELADEIANSKGALALEEKR--RLESRIAQLEEELEEEQGN 1744  
Qy 661 CQLAKKKQSETKLO-EELNKVLGKIKHFDPSK---AFHHEKSENFALKTPLKE 708  
Db 1745 TELYNDRLKKSILQIDQLTDLTAERSNAQKNENARQOOMDRQNKELTKLOE 1796

Search completed: December 20, 2002, 15:12:38  
Job time : 38.7824 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:06:33 ; Search time 37.4263 Seconds  
(without alignments)  
2581.255 Million cell updates/sec

Title: US-09-685-010-47

Perfect score: 3598

Sequence: 1 MSFPPKAPLRFNDPSCAPS.....LKEGNTNCRAPMECOESWK 725

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3598	100.0	725	AAW39165	Human RHAMM protei
2	3598	100.0	725	ABG50842	Human receptor for
3	3581	99.5	725	AAU11436	Human hyaluronic a
4	2670.5	74.2	713	AAU11438	Rat hyaluronic aci
5	2623	72.9	794	23 AAU11437	Mouse hyaluronic a
6	1953.5	54.3	630	18 AAW39166	Mouse RHAMM protei
7	1953	54.3	631	23 ABG50843	Mouse receptor for
8	1952	54.3	631	17 AAR99673	RHAMM 1-2a isoform
9	1819.5	50.6	606	17 AAR99673	Receptor for hyalu
10	1694	47.1	351	17 AAW01052	Human umbilical ve

11	1339	37.2	476	14	AAW39165	Hyaluronan recepto
12	429.5	11.9	1411	17	AAW02258	Nucleolar/endosoma
13	414	11.5	1988	22	AAW40999	Human polypeptide
14	414	11.5	1988	22	AAW41000	Human polypeptide
15	413	11.5	1388	23	AAU79590	Human kinesin moto
16	413	11.5	1388	23	AAE14400	Human kinesin supe
17	409.5	11.4	1948	22	ABG21233	Novel human diagno
18	407	11.3	2056	22	ABG59344	Drosophila melanog
19	406.5	11.3	1690	22	ABG61144	Drosophila melanog
20	406.5	11.3	1690	22	ABG61173	Human protein SEQ
21	404.5	11.2	1960	22	AAW78854	Novel human diagno
22	404.5	11.2	2143	22	ABG01716	Protein MYH11 diff
23	398.5	11.1	1857	23	AAU84350	Smooth muscle myos
24	398	11.1	1972	17	AAW00024	Human protein SEQ
25	397	11.0	1963	22	AAW79838	Novel human secret
26	396.5	11.0	2117	22	AAW32040	LexA/NumA fusion p
27	396.5	11.0	2192	18	AAW21732	GALA/HA/NumA fusio
28	396.5	11.0	2272	18	AAW21731	Sequence of the in
29	395.5	11.0	2101	15	AAW47173	Nuclear mitotic ap
30	395.5	11.0	2101	22	AAW65799	Novel human secret
31	395.5	11.0	2207	22	AAU32041	Human 160kD mediat
32	394.5	11.0	1427	12	AAW10534	Human NumA protein
33	393.5	10.9	2115	21	AAW49937	Human tumour marke
34	393	10.9	795	23	ABW77430	Human SCP-1 muten
35	393	10.9	976	22	AAW66581	Restin protein seq
36	392	10.9	1392	20	AAW06999	Novel human diagno
37	391.5	10.9	2246	22	ABG05850	Drosophila melanog
38	391	10.9	1489	22	ABW59948	Human RAD50. Homo
39	390.5	10.9	1312	18	AAW22775	Human homologue of
40	390.5	10.9	1312	19	AAW71295	Human alpha-myosin
41	389.5	10.8	2101	21	AAW49936	Human alpha-myosin
42	386	10.7	1939	23	ABW77096	Human protein sequ
43	381.5	10.6	1879	22	AAW25750	Human mitotin. Ho
44	381.5	10.6	2482	16	AAW72826	Human mitotin. Ho
45	381.5	10.6	2482	19	AAW23996	Human mitotin. Ho

ALIGNMENTS

RESULT 1

AAW39165

ID AAW39165 standard; Protein: 725 AA.

XX AC AAW39165;

XX DT 27-APR-1998 (first entry)

XX DE Human RHAMM protein.

XX KW Hyaluronan receptor; receptor for hyaluronic acid mediated motility;  
XX KW RHAMM; glycosaminoglycan; binding domain; human; oncogene; treatment;  
XX KW growth factor; cell locomotion disorder; dementia; detection;  
XX KW inflammatory disorder; autoimmune disease; diagnosis; prognosis.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT Region 478..480

XX FT Region /note= "Region indicated in specification"

XX FT Region 568..570

XX FT Region /note= "Region indicated in specification"

XX FT Region 589..591

XX FT Region /note= "Region indicated in specification"

XX FT Binding-site 636..646

XX FT Binding-site /note= "Hyaluronic acid binding domain"

XX FT Binding-site 658..667

XX FT Binding-site /note= "Hyaluronic binding domain"

XX PN WO9738098-A1.

XX PD 16-OCT-1997.

XX XX

PF 10-APR-1997; 97WO-CA00240.  
XX  
PR 10-APR-1996; 96GB-0007441.  
XX  
PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.  
PA (UYMA-) UNIV MANITOBA.  
XX  
PI Entwistle J, Turley EA;  
XX  
DR WPI; 1997-512715/47.  
DR N-PSDB; AAV02800.  
XX  
XX Isolated human receptor for hyaluronic acid mediated motility - used  
PT to develop products for treating e.g. tumours, inflammatory  
PT disorders, dementia, AIDS, diabetes and auto-immune diseases  
XX  
PS Claim 16; Page 46; 66pp; English.  
XX  
CC This sequence represents the human hyaluronan receptor which is also  
CC known as the receptor for hyaluronic acid mediated motility (RHAMM).  
CC Hyaluronan is a large glycosaminoglycan that is ubiquitous in the  
CC extracellular matrix and whose synthesis has been linked to cell  
CC migration, growth and transformation. It interacts with cell surfaces via  
CC specific protein receptors, e.g. RHAMM, that mediate many biological  
CC effects. The RHAMM/hyaluronic acid interaction is involved in  
CC oncogene- and growth factor-mediated cell locomotion. The products can be  
CC used in the treatment of disorders involving cell locomotion, e.g. tumour  
CC invasion, birth defects, acute and chronic inflammatory disorders,  
CC Alzheimer's and other forms of dementia, including Parkinson's and  
CC Huntington's diseases, AIDS, diabetes, autoimmune diseases, corneal  
CC dysplasia and hypertrophies, burns, surgical incisions and adhesions,  
CC strokes and multiple sclerosis. They can also be used in e.g. CNS and  
CC spinal cord regeneration, contraception and in vitro fertilisation and  
CC embryo development. The products can also be used in detection, diagnosis  
CC and prognosis.  
XX  
SQ Sequence 725 AA;  
Query Match 100.0%; Score 3598; DB 18; Length 725;  
Best Local Similarity 100.0%; Pred. No. 5,3e-185;  
Matches 725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSFPAKPLKRFNDPGCAPSPGAYDVKTLEVLKGPVSPQSKQRFPQKESKONLVNDKDT 60  
DB 1 MSFPAKPLKRFNDPGCAPSPGAYDVKTLEVLKGPVSPQSKQRFPQKESKONLVNDKDT 60  
QY 61 TLPSARKVKSSKESKOKNDKDLKLEKEIRVLQERGAQDRRIODLETELEKWEARL 120  
DB 61 TLPSARKVKSSKESKOKNDKDLKLEKEIRVLQERGAQDRRIODLETELEKWEARL 120  
QY 121 NAALREKTSLSANNATLEKQLTELTRTNELLSKFSSENGKNRLISLSELMKLNKRET 180  
DB 121 NAALREKTSLSANNATLEKQLTELTRTNELLSKFSSENGKNRLISLSELMKLNKRET 180  
QY 181 KMRGMAKQEGMEMQLVTVQSRLESQKIAOLEGKLVSIIEKEIDEKSETEKLEYIEE 240  
DB 181 KMRGMAKQEGMEMQLVTVQSRLESQKIAOLEGKLVSIIEKEIDEKSETEKLEYIEE 240  
QY 241 ISCASDQVEKYKLDIAOLEENLKENDELISLKSQLEENIVTLKQVEDLVNKCCLLEKE 300  
DB 241 ISCASDQVEKYKLDIAOLEENLKENDELISLKSQLEENIVTLKQVEDLVNKCCLLEKE 300  
QY 301 KEDVHNRNREHNENLNAEQNLKQFLEQOEHEKLOQKELQIDSLLOQEKELSSSLHQK 360  
DB 301 KEDVHNRNREHNENLNAEQNLKQFLEQOEHEKLOQKELQIDSLLOQEKELSSSLHQK 360  
QY 361 LCSFOEVMYKKNLFEELKQTLDELDKLQOKEQAEERLVKOLEEAKSRAEELKLEEK 420  
DB 361 LCSFOEVMYKKNLFEELKQTLDELDKLQOKEQAEERLVKOLEEAKSRAEELKLEEK 420  
QY 421 LKGRKAELEKSSAAQTATLLLOEKYDSMVOSLEDVTAQFESYKALTASETEDLKLENS 480  
DB 421 LKGRKAELEKSSAAQTATLLLOEKYDSMVOSLEDVTAQFESYKALTASETEDLKLENS 480

QY 481 LOEKAAGKAGKNAEDVQHOILATESSENOEYVRMLLDLQTSALKETEIKEITVSFLQKITD 540  
DB 481 LOEKAAGKAGKNAEDVQHOILATESSENOEYVRMLLDLQTSALKETEIKEITVSFLQKITD 540  
QY 541 LONQLKQOEEDPRKQLEDEEGRAKAKENTTABELTEINKWRLLYBELYNKTPFQQLDA 600  
DB 541 LONQLKQOEEDPRKQLEDEEGRAKAKENTTABELTEINKWRLLYBELYNKTPFQQLDA 600  
QY 601 FEVEKQALLNEHGAQEQOLNKIRDSYAKLLGHQNLKQIKKHVVKLKDNSQKSEVSKLR 660  
DB 601 FEVEKQALLNEHGAQEQOLNKIRDSYAKLLGHQNLKQIKKHVVKLKDNSQKSEVSKLR 660  
QY 661 COLAKKQKQSETKLQELNKLKLGKIKHFDPSKAPHHESKENFALKTPKLGKNTNCRYAPMBC 720  
DB 661 COLAKKQKQSETKLQELNKLKLGKIKHFDPSKAPHHESKENFALKTPKLGKNTNCRYAPMBC 720  
QY 721 QESWK 725  
DB 721 QESWK 725  
RESULT 2  
ABG60842  
ID ABG60842 standard; Protein; 725 AA.  
XX  
AC ABG60842;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
XX Human receptor for hyaluronan-mediated motility (RHAMM).  
XX  
XX Tissue disorder; response-to-injury process; cell proliferating;  
KW hyaluronic acid; HA; receptor for hyaluronan-mediated motility;  
KW RHAMM; inflammatory neurological disorder; Parkinson's disease;  
KW Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;  
KW inflammatory dermatosis; psoriasis; kidney fibrosis; inflammatory lung disease;  
KW stenosis; restenosis; cancer; kidney fibrosis; obesity; related disease;  
KW emphysema; asthma; cystic fibrosis; obesity; obesity related disease;  
KW lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;  
KW tissue transplantation; stroke; inflammatory response; fibrotic response;  
KW medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;  
KW myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;  
XX septic shock; thyroiditis; retinopathy.  
OS Homo sapiens.  
XX  
XX WO200228415-A1.  
XX  
XX 11-APR-2002.  
XX  
XX 05-OCT-2000; 2000WO-IB01534.  
XX  
XX 05-OCT-2000; 2000WO-IB01534.  
XX  
XX (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.  
XX  
XX Turley EA, Cruz TF;  
XX  
XX WPI; 2002-435298/46.  
XX  
XX Treating tissue disorder associated with response-to-injury process or  
PT proliferating cells in mammals, e.g. fibrosis, inflammation, by  
PT administering a compound that alters activity of transition molecules  
PT within a cell -  
XX  
XX Example 30; Fig 50; 215pp; English.  
XX  
XX The invention describes a method of treating a tissue disorder associated  
CC with response-to-injury process or proliferating cells in a patient,  
CC comprising administering a polypeptide (I) which binds hyaluronic acid  
CC (HA), an antibody which binds one of domains D1-D5 of Receptor for  
CC hyaluronan-mediated motility (RHAMM), a polypeptide fragment encoding

CC any of D1-D5 of RHAMM, or a vector which expresses antisense RHAMM,  
CC antibodies or a polypeptide fragment. The method is useful for treating a  
CC patient with an inflammatory neurological disorder such as Parkinson's  
CC disease, Alzheimer's disease, arthritis including rheumatoid arthritis,  
CC osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),  
CC inflammatory bowel disease, stenosis or restenosis, cancer, kidney  
CC fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic  
CC fibrosis), obesity or obesity related diseases, lupus, cardiovascular  
CC disease (e.g. atherosclerosis), and wound especially surgical excision  
CC adhesions, to prevent scar and also for treating or preventing diabetes  
CC mellitus. The method is also useful for treating tissue transplantation  
CC (e.g. skin grafts), stroke, inflammatory responses or fibrotic response  
CC associated with medical implants such as hip implants, vascular wraps and  
CC catheters), inflammatory diseases such as AIDS, myocardial and hepatic  
CC fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,  
CC hepatitis, septic shock, thyroiditis, and retinopathy. This sequence  
CC represents a receptor for hyaluronan-mediated motility protein used in  
CC the method of treating a tissue disorder described in the invention.  
XX  
SQ

Sequence 725 AA;

Query Match 100.0%; Score 3598; DB 23; Length 725;  
Best Local Similarity 100.0%; Pred. No. 5.3e-185;  
Matches 725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFPKAPLKRENDPGCAPSPGAYDVKTLEVLKGPVFSQKSFQKESQKQNLNVKDT 60  
Db 1 MSFPKAPLKRENDPGCAPSPGAYDVKTLEVLKGPVFSQKSFQKESQKQNLNVKDT 60  
Qy 61 TLPASARKVKSSESQKNDKDLKILEKEIRVLLQGERGAQRRIQDLETELEKMEARL 120  
Db 61 TLPASARKVKSSESQKNDKDLKILEKEIRVLLQGERGAQRRIQDLETELEKMEARL 120  
Qy 121 NAALREKTSLSANNATLEKQIETRLNELLKSFSENGNOKNRLILSLMLKLNRET 180  
Db 121 NAALREKTSLSANNATLEKQIETRLNELLKSFSENGNOKNRLILSLMLKLNRET 180  
Qy 181 KMRGMAKQEGEMKLVQTRSLSESGKIAQLGKGLVSIKEKIDSEKSEKLVETEE 240  
Db 181 KMRGMAKQEGEMKLVQTRSLSESGKIAQLGKGLVSIKEKIDSEKSEKLVETEE 240  
Qy 241 ISCASDQVEKYKLDIALEENLKEKNDIISLKOSLEENIVILSKQVEDLVNKKOLEKE 300  
Db 241 ISCASDQVEKYKLDIALEENLKEKNDIISLKOSLEENIVILSKQVEDLVNKKOLEKE 300  
Qy 301 KEDHVNRRHNEHNLNEMQNLKQFVLEQOEHEKLOQKELQIDSLLOQEKELSSSLHQK 360  
Db 301 KEDHVNRRHNEHNLNEMQNLKQFVLEQOEHEKLOQKELQIDSLLOQEKELSSSLHQK 360  
Qy 361 LCSFQEEVMVKEKNLFEBELKQTLDELQKQEQAEERLVKQLEEEAKSRAEELKLLLEK 420  
Db 361 LCSFQEEVMVKEKNLFEBELKQTLDELQKQEQAEERLVKQLEEEAKSRAEELKLLLEK 420  
Qy 421 LKGEAELEKSSAAHTQATLLQEKYSWQSLVEDVTAQFYSYKALTASEIEDLKLENS 480  
Db 421 LKGEAELEKSSAAHTQATLLQEKYSWQSLVEDVTAQFYSYKALTASEIEDLKLENS 480  
Qy 481 LQEAAGKAGNAEDVOHILATESNQEVVRLDLQTKSALKETEKEIVSVFLQKITD 540  
Db 481 LQEAAGKAGNAEDVOHILATESNQEVVRLDLQTKSALKETEKEIVSVFLQKITD 540  
Qy 541 LQNLKQOEEDFRKQLEDEGRKAEKENTTAELTEENKRWLLYBELYNKTKPQLOLDA 600  
Db 541 LQNLKQOEEDFRKQLEDEGRKAEKENTTAELTEENKRWLLYBELYNKTKPQLOLDA 600  
Qy 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQIKIHVVKLDKENSQKSEVSKLR 660  
Db 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQIKIHVVKLDKENSQKSEVSKLR 660  
Qy 661 QOLAKKKQSETKLOEELNKVLGKIHFDPSKAFHHESKENFALKTPLEKGNVNCYRAPMEC 720  
Db 661 QOLAKKKQSETKLOEELNKVLGKIHFDPSKAFHHESKENFALKTPLEKGNVNCYRAPMEC 720

Qy 721 QESWK 725  
Db 721 QESWK 725  
RESULT 3  
ID AAU11436 standard; Protein; 725 AA.  
XX AAU11436;  
XX 12-MAR-2002 (first entry)  
XX Human hyaluronic acid binding protein RHAMM.  
XX Human; hyaluronic acid binding protein; RHAMM; gene therapy;  
KW receptor for HA mediated mobility; immunosuppressive; cytostatic.  
KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FT Misc-difference 278 /note= "Encoded by GAA"  
FT Misc-difference 299 /note= "Encoded by AAA"  
FT Misc-difference 323 /note= "Encoded by AAA"  
FT Misc-difference 331 /note= "Encoded by CAG"  
XX WO200180899-A2.  
XX 01-NOV-2001.  
XX 20-APR-2001; 2001WO-CA00533.  
XX 20-APR-2000; 2000US-198613P.  
XX (CANG-) CANGENE CORP.  
XX Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;  
XX WPI: 2002-075094/10.  
XX N-PSDB; AAS17496.  
XX Protein conjugates that selectively target certain tissues and organs  
XX useful for treating and preventing various diseases, comprises  
XX glucose-aminoglycan-targeting domain conjugated to a therapeutic  
XX protein .  
XX Claim 6; Page 115; 121pp; English.  
XX The invention relates to a conjugate comprising an hyaluronic acid (HA)  
XX -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptidic  
XX contiguous with, or coupled to a polypeptide conjugated to a therapeutic  
XX agent, and the polynucleotides encoding them. Also included is a method  
XX for preparation of the HA-binding protein by inserting a first nucleotide  
XX sequence encoding a HA-binding protein directly linked to a second  
XX nucleotide sequence encoding a therapeutic protein into a suitable  
XX vector, expressing the vector in an acceptable host, purifying conjugate  
XX molecule from host or expression medium. The composition is useful for  
XX altering in vivo the distribution of a therapeutic agent comprising  
XX administering the composition to the animal where conjugate molecule will  
XX distribute primarily in tissues and organs containing high levels of  
XX endogenous HA and for treating mammal with a disorder where a diseased  
XX tissue of the mammal contains high level of HA e.g. rheumatoid  
XX arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic  
XX dosages required also translates into lower immunogenicity of the  
XX conjugated protein as compared to the native protein. As a result,  
XX conjugates improve patient compliance and reduce direct and indirect  
XX costs associated with the drug substance and its administration.  
XX Conjugates allows for the use, where appropriate, of lower, safer,

CC dosages as compared to the conventional dosage requirements for the  
CC unconjugated corresponding therapeutic agent. Conjugate molecules have an  
CC increased half-life and potency, resulting in prolonged circulation of  
CC the molecule, efficient distribution into the target tissues, and  
CC increased bioavailability. The present sequence represents a RHAMM  
CC protein.

XX SQ Sequence 725 AA;

Query Match 99.5%; Score 3581; DB 23; Length 725;  
Best Local Similarity 99.4%; Pred. No. 4.3e-184;  
Matches 721; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSFPKAPLKRENDPSGCAPSGAYDVKTLEVLKGPVSPQSKQRFKQKESQNLNVKDT 60  
Db 1 MSFPKAPLKRENDPSGCAPSGAYDVKTLEVLKGPVSPQSKQRFKQKESQNLNVKDT 60  
Qy 61 TLPASARKVSKESKQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETEKMEARL 120  
Db 61 TLPASARKVSKESKQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETEKMEARL 120  
Qy 121 NAALREKTSLSANNATLEKQIETRTNELLKSKFSENGKNLRIILSLKMLKRNKRET 180  
Db 121 NAALREKTSLSANNATLEKQIETRTNELLKSKFSENGKNLRIILSLKMLKRNKRET 180  
Qy 181 KMRGMAKQEGMEMKQVQTSLSQKIAQLEGKLVSIKIEKIDKSEKLEVEE 240  
Db 181 KMRGMAKQEGMEMKQVQTSLSQKIAQLEGKLVSIKIEKIDKSEKLEVEE 240  
Qy 241 ISCASQVQKYLDTAILEENKEKNDLSLKQSLNIVILSKQVEDLVNKKOLEKE 300  
Db 241 ISCASQVQKYLDTAILEENKEKNDLSLKQSLNIVILSKQVEDLVNKKOLEKE 300  
Qy 301 KEDHVRNREHNENLNAEMONLKQFLEQOEHEKLOQKELQIDSLQOEKELSSLSHQK 360  
Db 301 KEDHVRNREHNENLNAEMONLKQFLEQOEHEKLOQKELQIDSLQOEKELSSLSHQK 360  
Qy 361 LCSFOEMVKEKNLFEELKQTLDELQKQEEQAEVLVQLEBEAKSRAEELKLEEK 420  
Db 361 LCSFOEMVKEKNLFEELKQTLDELQKQEEQAEVLVQLEBEAKSRAEELKLEEK 420  
Qy 421 LKGAELKSSAAHTQTLILLOEKYDSMWQSLQEDVTAQFESYKALTASEIDKLKENS 480  
Db 421 LKGAELKSSAAHTQTLILLOEKYDSMWQSLQEDVTAQFESYKALTASEIDKLKENS 480  
Qy 481 LQEKAAKAGNAEDVQHOILATESNQEVVRLMLDLQTKSALKEIETVSLFKITD 540  
Db 481 LQEKAAKAGNAEDVQHOILATESNQEVVRLMLDLQTKSALKEIETVSLFKITD 540  
Qy 541 LQNLQKQEDDFRQLEDEGRKAENKNTAELTEENKWRLLYBEELYNKTKPFQLQDA 600  
Db 541 LQNLQKQEDDFRQLEDEGRKAENKNTAELTEENKWRLLYBEELYNKTKPFQLQDA 600  
Qy 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLGHQNLKQIKHVKVLKDNSQKSEVSKLR 660  
Db 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLGHQNLKQIKHVKVLKDNSQKSEVSKLR 660  
Qy 661 COLAKKQKSETKLOELKNVILGKHFDPSKAFHHESKENFALKTPKKEGNTNCRAPMEC 720  
Db 661 COLAKKQKSETKLOELKNVILGKHFDPSKAFHHESKENFALKTPKKEGNTNCRAPMEC 720  
Qy 721 QESWK 725  
Db 721 QESWK 725

RESULT 4  
AAU11438  
ID AAU11438 standard; Protein: 713 AA.  
XX  
AC  
XX AAU11438;  
DT 12-MAR-2002 (first entry)

XX DE Rat hyaluronic acid binding protein RHAMM.  
XX KW Rat; hyaluronic acid binding protein; RHAMM; gene therapy;  
KW receptor for HA mediated mobility; immunosuppressive; cytostatic.  
KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;  
XX OS Rattus norvegicus.  
XX PN WO200180899-A2.  
XX XX  
PD 01-NOV-2001.  
XX 20-APR-2001; 2001WO-CA00533.  
XX PF  
XX 20-APR-2000; 2000US-198613P.  
XX PR  
XX (CANG-) CANGENE CORP.  
XX PA  
XX Wołoski BMR, Williams AM, Sereda TJ, Wiebe DJ;  
PI  
XX WPI: 2002-075094/10.  
XX DR N-PSDB; AAS17498.  
XX DR  
XX Protein conjugates that selectively target certain tissues and organs  
PT useful for treating and preventing various diseases, comprises  
PT glucose-aminoglycan-targeting domain conjugated to a therapeutic  
PT protein -  
XX  
PS Claim 6; Page 116-117; 121pp; English.

CC The invention relates to a conjugate comprising an hyaluronic acid (HA)  
CC -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide  
CC contiguous with, or coupled to a polypeptide conjugated to a therapeutic  
CC agent, and the polynucleotides encoding them. Also included is a method  
CC for preparation of the HA-binding protein by inserting a first nucleotide  
CC sequence encoding a HA-binding protein directly linked to a second  
CC nucleotide sequence encoding a therapeutic protein into a suitable  
CC vector, expressing the vector in an acceptable host, purifying conjugate  
CC molecule from host or expression medium. The composition is useful for  
CC altering in vivo the distribution of a therapeutic agent comprising  
CC administering the composition to the animal where conjugate molecule will  
CC distribute primarily in tissues and organs containing high levels of  
CC endogenous HA and for treating mammal with a disorder where a diseased  
CC tissue of the mammal contains high level of HA e.g. rheumatoid  
CC arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic  
CC dosages required also translates into lower immunogenicity of the  
CC conjugated protein as compared to the native protein. As a result,  
CC costs associated with the drug substance and its administration.  
CC Conjugates allows for the use, where appropriate, of lower, safer,  
CC dosages as compared to the conventional dosage requirements for the  
CC unconjugated corresponding therapeutic agent. Conjugate molecules have an  
CC increased half-life and potency, resulting in prolonged circulation of  
CC the molecule, efficient distribution into the target tissues, and  
CC increased bioavailability. The present sequence represents a RHAMM  
CC protein.

XX SQ Sequence 713 AA;

Query Match 74.2%; Score 2670.5; DB 23; Length 713;  
Best Local Similarity 75.3%; Pred. No. 2.6e-135;  
Matches 537; Conservative 71; Mismatches 104; Indels 1; Gaps 1;

Qy 1 MSFPKAPLKRENDPSGCAPSGAYDVKTLEVLKGPVSPQSKQRFKQKESQNLNVKDT 60  
Db 1 MSFPKAPLKRENDPSGCAPSGAYDVKTSESTKGPVSPQSKQRFKQKESQNLNVKDT 60  
Qy 61 TLPASARKVSKESKQKNDKDLKILEKEIRVLLQERGAQDRRIQDLETEKMEARL 120  
Db 61 TLLASAKAKTIVSKESQKNDKVKRLEKEIHVLLQERGTQDKRQIQMESELENTEANL 120  
Qy 121 NAALREKTSLSANNATLEKQIETRTNELLKSKFSENGKNLRIILSLKMLKRNKRET 180

Db 121 NAPVTEKPSLANNASLEKRLTELTRANELKSKFSEDAHQNMRTLSLEMLKRNKRET 180  
Qy 181 KMRGMAKQEGMEKQVOTORSLEESQKIAQLEGKLVSIIEKIDSEKSETEKLEYIEE 240  
Db 181 KMRGMAKQEGMEKQVOTORSLEESQKIAQLEGKLVSIIEKIDSEKSETEKLEYIEE 240  
Qy 241 ISCASDQVEKYKLDIAQLEENLKEKNDLILSKQSLSEENIVILSKQVEDLVNKKQLEKE 300  
Db 241 ISCASDQVEKYKLDIAQLEENLKEKNDLILSKQSLSEENIVILSKQVEDLVNKKQLEKE 299  
Qy 301 KEDHVNRRNHNENLNAMQNLKQKFILEQOEHEKLOKQELQIDSLSQOEKELSSSLHQK 360  
Db 300 RDDLVSKDRRAESLSAEMQVLTETKLLEROEYKLOQNELQSSLLQOEKELSAHQOQ 359  
Qy 361 LCSQOEEMVKELFEPELKOTLDELKLOQEKQOAEPLVKQLEEAESAEELKLEEK 420  
Db 360 LCSQOEEMTSERNVFEKQKLALDELDAVQKQEKQSEKLVKQLEBEYKSTAEQURRDDL 419  
Qy 421 LKGEAELEKSSAHTQATLLQKQYDSMWQSLQEDVTAQFESYKALTAASEIEDLKLENS 480  
Db 420 LREKEIELEKRTAAHAQVIAQEKYSDTAQTLDVTAQLESYKSSSTLKEIEDLKLENI 479  
Qy 481 LOEAAKAGKNAEDVQOILATESNQBYVRMLDLQTKSALKETEKEITVSVLOKTD 540  
Db 480 LOEKVMAEKREVDVQOILTAESTNQBYAKVQDQNSSTLKAEKEITVSVLEKTD 539  
Qy 541 LQNLKQOEEDFRKOLEDEGRKAENKNTAELTEENKRWLLYEELNKKTKPQQLQDA 600  
Db 540 LQNLQROQNEDFRQLEEGAKMTEKETAVTELTMEINRWLLYEELFDKTKPQQQLDA 599  
Qy 601 FEVEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLDKDNSOLKSEVSKLR 660  
Db 600 FEAEKQALLNEHGAQEQNLKIRDSYAKLLGHQNLKQKIKHVVKLDKDNSOLKSEVSKLR 659  
Qy 661 COLAKKQOSETKLOELNKNVIGIKHFDPSKAFHSHESKENFALKPLKEGNTNC 713  
Db 660 SOLAKRKQNELRQELGDKALGIRHFDPSKAFCHESKENVTLPKLEGPNPC 712

RESULT 5  
AAU11437  
ID AAU11437 standard; Protein; 794 AA.  
XX AC AAU11437;  
XX DT 12-MAR-2002 (first-entry)  
XX DE Mouse hyaluronic acid binding protein RHAMM.  
XX KW Mouse; hyaluronic acid binding protein; RHAMM; gene therapy;  
KW KW receptor for HA mediated mobility; immunosuppressive; cytostatic.  
KW KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;  
XX OS Mus musculus.  
XX FH Key Location/Qualifiers  
FT FT Misc-difference 55 /note= "Encoded by AAC"  
FT FT Misc-difference 71 /note= "Encoded by ACG"  
FT FT Misc-difference 89 /note= "Encoded by CAA"  
FT FT Misc-difference 91 /note= "Encoded by CAC"  
FT FT Misc-difference 540 /note= "Encoded by ACT"  
FT FT Misc-difference 668 /note= "Encoded by GAT"  
XX FN W0200180899-A2.  
XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-CA00533.  
XX PR 20-APR-2000; 2000US-198613P.  
XX PA (CANG-) CANGENE CORP.  
XX PI Wołoski BMR, Williams AM, Sereda TJ, Wiebe DJ;  
XX PI WPI; 2002-075094/10.  
XX DR N-PSDB; AAS17497.  
XX PT Protein conjugates that selectively target certain tissues and organs  
PT useful for treating and preventing various diseases, comprises  
PT glucose-aminoglycan-targeting domain conjugated to a therapeutic  
PT protein  
XX PS Claim 6; Page 116; 121pp; English.  
XX CC The invention relates to a conjugate comprising an hyaluronic acid (HA)  
CC -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide  
CC contiguous with, or coupled to a polypeptide conjugated to a therapeutic  
CC agent, and the polynucleotides encoding them. Also included is a method  
CC for preparation of the HA-binding protein by inserting a first nucleotide  
CC sequence encoding a HA-binding protein directly linked to a second  
CC nucleotide sequence encoding a therapeutic protein into a suitable  
CC vector, expressing the vector in an acceptable host, purifying conjugate  
CC molecule from host or expression medium. The composition is useful for  
CC altering in vivo the distribution of a therapeutic agent comprising  
CC administering the composition to the animal where conjugate molecule will  
CC distribute primarily in tissues and organs containing high levels of  
CC endogenous HA and for treating mammal with a disorder where a diseased  
CC tissue of the mammal contains high level of HA e.g. rheumatoid  
CC arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic  
CC dosages required also translates into lower immunogenicity of the  
CC conjugated protein as compared to the native protein. As a result,  
CC conjugates improve patient compliance and reduce direct and indirect  
CC costs associated with the drug substance and its administration.  
CC Conjugates allows for the use, where appropriate, of lower, safer,  
CC dosages as compared to the conventional dosage requirements for the  
CC unconjugated corresponding therapeutic agent. Conjugate molecules have an  
CC increased half-life and potency, resulting in prolonged circulation of  
CC the molecule, efficient distribution into the target tissues, and  
CC increased bioavailability. The present sequence represents a RHAMM  
XX protein.  
XX SQ Sequence 794 AA;  
Query Match 72.9%; Score 2623; DB 23; Length 794;  
Best Local Similarity 68.0%; Pred. No. 1e-132;  
Matches 542; Conservative 67; Mismatches 100; Indels 88; Gaps 3;  
Qy 1 MSFPAKPLKRNNDPSGCAPSPGAYDVKYLEVLKVPVSFKSQSRFKQKESKQNLNVKDKT 60  
Db 1 MSFPAKPLKRNNDPSGCAPSPGAYDVKYLEVLKVPVSFKSQSRFKQKESKQNLNVKDKT 60  
Qy 61 TLPASARKVKSSEKSKQNDKDLKILEKEIRVLLQERGAODRRIDOLETELEKWEARL 120  
Db 61 TLLASAKKAKKSVSKDKQNDKQVRLKEIRALLQERGTODKRIQDMESLEKTEAKL 120  
Qy 121 NAALREKTSLSANNATLEKQIETLRTNELLKSKFSENGNKNLRTLSLEMLKRNKRET 180  
Db 121 NAAVREKTSLSASNASLEKRLTELTRANELKSKFSENGHGNKMSALSLEMLKRNKRET 180  
Qy 181 KMRGMAKQEGMEKQVOTORSLEESQKIAQLEGKLVSIIEKIDSEKSETEKLEYIEE 240  
Db 181 KMRSMVVKQEGMELKQATQDLTSEKGIQVLEGKLVSIIEKIDSEKSETEKLEYIQE 240  
Qy 241 ISCASDQVEKYKLDIAQLEENLKEKNDLILSKQSLSEENIVILSKQVEDLVNKKQLEKE 300  
Db 241 ISCASDQVEKCKVDIAQLEEDLKEDREILSLKQSLSEENIT-FSKQIEDLVTKCOLLETE 299  
Qy 301 KEDHVNRRNHNENLNAMQNLKQKFILEQOEHEKLOKQELQIDSLSQOEKELSSSLHQK 360





Qy 691 LGIKHFDPSKAFHSHESKENFALKTPKKEGNTNC 713  
 ID ABG60843  
 Db 600 LGIRHFDPSKAFCHASKENF---TPLKEGNPNC 629

RESULT 7  
 ID ABG60843 standard; Protein; 631 AA.

XX AC ABG60843;  
 XX 13-AUG-2002 (first entry)  
 XX Mouse receptor for hyaluronan-mediated motility (RHAMM).

XX Tissue disorder; response-to-injury process; cell proliferating;  
 KW hyaluronic acid; HA; receptor for hyaluronan-mediated motility;  
 KW RHAMM; inflammatory neurological disorder; Parkinson's disease;  
 KW Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;  
 KW inflammatory dermatosis; psoriasis; inflammatory bowel disease;  
 KW stenosis; restenosis; cancer; kidney fibrosis; inflammatory lung disease;  
 KW emphysema; asthma; cystic fibrosis; obesity; obesity related disease;  
 KW lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;  
 KW tissue transplantation; stroke; inflammatory response; fibrotic response;  
 KW medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;  
 KW myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;  
 KW septic shock; thyroiditis; retinopathy.

XX Mus musculus.  
 OS  
 XX WO200228415-A1.  
 XX 11-APR-2002.  
 XX 05-OCT-2000; 2000WO-IB01534.  
 XX 05-OCT-2000; 2000WO-IB01534.  
 XX (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.  
 XX Turley EA, Cruz TF;  
 XX WPI; 2002-435298/46.

XX Treating tissue disorder associated with response-to-injury process or  
 XX proliferating cells in mammals, e.g. fibrosis, inflammation, by  
 XX administering a compound that alters activity of transition molecules  
 XX within a cell

XX Disclosure; Fig 50; 215pp; English.

XX The invention describes a method of treating a tissue disorder associated  
 XX with response-to-injury process or proliferating cells in a patient,  
 XX comprising administering a polypeptide (I) which binds hyaluronic acid  
 XX (HA), an antibody which binds one of domains D1-D5 of Receptor for  
 XX hyaluronan-mediated motility (RHAMM), a polypeptide fragment encoding  
 XX any of D1-D5 of RHAMM, or a vector which expresses antisense RHAMM,  
 XX antibodies or a polypeptide fragment. The method is useful for treating a  
 XX patient with an inflammatory neurological disorder such as Parkinson's  
 XX disease, Alzheimer's disease, arthritis including rheumatoid arthritis,  
 XX osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),  
 XX inflammatory bowel disease, stenosis or restenosis, cancer, kidney  
 XX fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic  
 XX fibrosis), obesity or obesity related diseases, lupus, cardiovascular  
 XX disease (e.g. atherosclerosis), and wound especially surgical excision  
 XX adhesions, to prevent scar and also for treating or preventing diabetes  
 XX mellitus. The method is also useful for treating tissue transplantation  
 XX (e.g. skin grafts), stroke, inflammatory responses or fibrotic response  
 XX associated with medical implants such as hip implants, vascular wraps and  
 XX catheters), inflammatory diseases such as AIDS, myocardial and hepatic  
 XX fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,  
 XX hepatitis, septic shock, thyroiditis, and retinopathy. This sequence

CC represents a receptor for hyaluronan-mediated motility protein used in  
 CC the method of treating a tissue disorder described in the invention.

XX  
 SQ Sequence 631 AA;

Query Match 54.3%; Score 1953; DB 23; Length 631;  
 Best Local Similarity 64.8%; Pred. No. 6.2e-97;  
 Matches 411; Conservative 49; Mismatches 86; Indels 88; Gaps 3;

Qy 164 LRILSLMLKLRNKRRETKMRGMKAKOEGMEMKLVQVTSRLEESOGKIAQLEKGLVSIKRE 223  
 Db 1 MRALSLEMLKLRNKRRETKMRGMKAKOEGMEMKLVQVTSRLEESOGKIAQLEKGLVSIKRE 60  
 Qy 224 KIDEKSETEKLEIYEIEISCASDQVEKYKLDIALEENLKEKNDREILSKOSLEENIVIL 283  
 Db 61 KIDEKSETEKLEIYEIEISCASDQVEKYKLDIALEENLKEKNDREILSKOSLEENIT-F 119  
 Qy 284 SKQVEDLVNKKOLEKEKEDHVRNRHNNENINAMQNLKOKFILEQOEHEKLOKQELQI 343  
 Db 120 SKQIEDLVNKKOLEKEDHVRNRHNNENINAMQNLKOKFILEQOEHEKLOKQELQI 179  
 Qy 344 DSLQOEKELSSSLHOKLCSFOEEMVKEKLFEEELKOTLDLKLQOEKQOERLVLKOL 403  
 Db 180 QSLQOEKELSSSLHOKLCSFOEEMVKEKLFEEELKOTLDLKLQOEKQOERLVLKOL 239  
 Qy 404 EEEAKSRAEELKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLE 457  
 Db 240 EEEKSTAEQLTRLDNLRREKEVELEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLEKLE 299  
 Qy 458 ----- 457  
 Db 300 QEKYNDTAQSLRDVTAQLESEOEKYNDTAQSLRDVTAQLESEOEKYNDTAQSLRDVTAQ 359  
 Qy 458 -----AQESYKALTASEIEDLKLENSLOEKAAKAGNAEDVOHQI 499  
 Db 360 ESOEYKNDTAQSLRDVTAQLESEOEKYNDTAQSLRDVTAQLESEOEKYNDTAQSLRDV 419  
 Qy 500 LATESSNOEYVRLDLDLQTSALKETEIKETVSPLOKQITDLONOLKQOEEDFRKLEDE 559  
 Db 420 LTAESTNOEYARMVDLQNRSTLKEEKEIKETSSFEKLEITDLKNLRQDEDFRQLEEK 479  
 Qy 560 EGRKAKEKNTTAELEETINKWRLLYEELYNKTPFQLQDAFEVEKQALLNHEGAAQEO 619  
 Db 480 GKRTAEKENVMTETLMEINKWRLLYEELYEKTPFQQLDAFEVEKQALLNHEGATQEO 539  
 Qy 620 NKIRDSYAKILGHQNLKQIKHVKLKDENSOLKSEVSKLRCQLAKKQSEKLOEELNK 679  
 Db 540 NKIRDSYAKILGHQNLKQIKHVKLKDENSOLKSEVSKLRCQLAKKQSEKLOEELNK 679  
 Qy 680 VLGIKHFDPSKAFHSHESKENFALKTPKKEGNTNC 713  
 Db 600 ALGIRHFDPSKAFCHASKENF---TPLKEGNPNC 630

RESULT 8  
 AAR99675  
 ID AAR99675 standard; Protein; 631 AA.  
 XX AC AAR99675;  
 XX 10-OCT-1996 (first entry)  
 XX RHAMM 1-2a isoform.  
 XX RHAMM 1-2a; receptor for hyaluronic acid mediated motility;  
 KW hyaluronan receptor; cell locomotion; cell proliferation;  
 KW breast cancer; therapy.  
 XX Mus sp.  
 XX Key  
 XX Region 55..79  
 XX /note= "exon 2A-encoded region"







Matches 183; Conservative 146; Mismatches 257; Indels 141; Gaps 31;

QY	30	EVLKGP----	VSFQKSQRFKQESKQNINVDKDTTLPASARKVSVSESKSQKNRDXDL	85
Db	275	ELAKGQFVAVVVOELQKLK----	SSVNELTKQNKOTLTENLLKKEDDYTLKEEKHNEE---	328
QY	86	KILEIKERVILLBERGAADRRIQDLETELEKMEARLNAALUREKTSLSANNATLEKOLIELT	145	
Db	329	SYSKKNIQTALHQK-----	DLDCCQLQSRLSAS---ETSLH---RHVELSEKG	371
QY	146	RTNELLKSPFSENGNKNLRILISLEMLKURNRETK-MRGMMAKOE--GMEMKLVQTVORS	202	
Db	372	EATOKLKEELSE--VETKYOHLKAEPKOLQQOREEQHGLOQSEINOHLSHKLLETERQ	429	
QY	203	LBSQCK-----	IALEGKLVSI---KKIDEKSTEXLEYIEB-	240
Db	430	LGEAHRLEKEQRQLSESSEKLMXDEQQVADQLKLSLEQEOLKEKVNTNSTLQHOLDKTQK	489	
QY	241	-----ISCASDOVBKYKLDIAQLEENLEKENDETSLUK--QSLSEENVITLSKOVEDL	290	
Db	490	HQBQQAALQOSTTAKLREAANDLEQVLRIGDKDDQKIQNLEALLQKSKEWISLUKEREDL	549	
QY	291	NVKCQLLEKEKEEDHVNRNHNENNAEMONLKORFIPEQBHEKLIQQKELAIDSLLQOE	350	
Db	550	YAKIQAGEGETA-VLNQOLEKNHTLQEQVTQLEKLNQSESHKQAE--NLHDQVQEQ	605	
QY	351	KELSSLSHOKLCSFOEMYKKNLFEEBELKQTDDELDKLQKQEBQAEURLVKOLEBEAKSR	410	
Db	606	KAHLRAAQDRVLSSL-ETSYNELNSQLNESKEKVSQILD--IQIKAKTELIVSAEAATAQR	662	
QY	411	AE---ELKULEELEKKGAELEKSSAAHTQATLLIQQEKYDSMVOSILEDYTAOFESYKALT	467	
Db	663	ADLQNHILDTAQNALQDQOGBELKNITQLDQVTAKLQDKQEHCSQLESHELUKEYKEYLSL-	721	
QY	468	ASBIEDL-----KLENSSUQEKAAGKNAEDVQHO-----ILATESNQ-----EY	509	
Db	722	EQKTEELGGIKKLEADSLEVRSKAE-QALQDLQOQRQLNTDLELATELSQKLEMEKEI	780	
QY	510	VRML-LDLQTKSALKETEKETVSPLOKITDLQNLKOOBEDFRKQLEDSEGKRAEKN	568	
Db	781	VSTRULDQKKS-----EALESIHKOLT-----KOEFE---KQILKQDFETLSQFT	823	
QY	569	TTAELTFEINKWRLLYEELNYNTPKPQLDAFEVEKQOALLNEHGAQAQOLANKIRDSY--	626	
Db	824	-----KIQHEELNNRIQTTVTLEQVKVMEKEALMTLETSTVDKLSKVSDSLKN	871	
QY	627	-----AKULLGHONLKQIKKHVVVLKLDENSEQLSKSEVSKLRCOLAKKKQSETKL	673	
Db	872	SKSEFEKENQGKAAILDLEKTCCKELKHQLOVOMENT-LK-EQKELKLSLEKEKEASHOL	929	
QY	674	QEELNVK	680	
Db	930	KLEINSM	936	

RESULT 13  
AAM40999

ID	AAM40999 standard; Protein; 1988 AA.
1	100
2	100
3	100
4	100
5	100
6	100
7	100
8	100
9	100
10	100
11	100
12	100
13	100
14	100
15	100
16	100
17	100
18	100
19	100
20	100
21	100
22	100
23	100
24	100
25	100
26	100
27	100
28	100
29	100
30	100
31	100
32	100
33	100
34	100
35	100
36	100
37	100
38	100
39	100
40	100
41	100
42	100
43	100
44	100
45	100
46	100
47	100
48	100
49	100
50	100
51	100
52	100
53	100
54	100
55	100
56	100
57	100
58	100
59	100
60	100
61	100
62	100
63	100
64	100
65	100
66	100
67	100
68	100
69	100
70	100
71	100
72	100
73	100
74	100
75	100
76	100
77	100
78	100
79	100
80	100
81	100
82	100
83	100
84	100
85	100
86	100
87	100
88	100
89	100
90	100
91	100
92	100
93	100
94	100
95	100
96	100
97	100
98	100
99	100
100	100

AC AAM40999;

DT 22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 5930

Human: nootropic: immunosuppressant

KW peripneural nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Sly-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX Homo sapiens.

5

XX PN XX PD XX PF XX PR PR PR PR PR PR PR PA XX XI PI PI PI XX DR DR XX PT PT XX PS XX CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC

Query Match	11.5%;	Score 414;	DB 22;	Length 1988;
Best Local Similarity	23.4%;	Pred. No. 6.2e-14;		
Matches 199;	Conservative 150;	Mismatches 310;	Indels 190;	Gaps 29;

Qy	32	LKGPVFSQSRFKQOKESQNVLNVDKDTTL--PASARKV-----KSS	72
		: : :     : : :     : : :     : : :     : : :     : : :	
Db	1086	LQAIDELQLAQAKEEGLALRGDDTFLHKNALKVRLQIAIELQEDFESERAS	1145
		: : :     : : :     : : :     : : :     : : :     : : :	
Qy	73	ESKSESOKND--KDKILKEKRVLLQERGAQ-----DRR	105
		: :     : : :     : : :     : : :     : : :     : : :	
Db	1146	RNAEKQRDLSEEBALYATELEDLTDTAAOQELRTKREQVEAEKKALEETKNHEAQ	1205
		: : :     : : :     : : :     : : :     : : :     : : :	
Qy	106	IDOLE----TELEKMBARLNAALREKTSUSANNATLEKOLIBELTNELLNKSFKSNGNQ	161
		: : :     : : :     : : :     : : :     : : :     : : :	
Db	1206	IDMRQRHATALESELSEQEAQRKANLEKNKGLETDNKELACEVKLQQVKAESHHK	1265
		: : :     : : :     : : :     : : :     : : :     : : :	
Qy	162	K-----NLRI--LSLELMKLNRK-----RFTKMGRMAKOE--GM	192
		: : :     : : :     : : :     : : :     : : :     : : :	
Db	1266	RKKLDQAQVELHAKVSEGDRLRVLAEKAKLQNELDNVSTILLEBAEKKGIKFAKDAASL	1325
		: : :     : : :     : : :     : : :     : : :     : : :	
Qy	193	EMKLQVTQRSLESQGIKIQLEGKLVSIPEKIDEKSETFEKLLIYEISCASDQVEYK	252
		: : :     : : :     : : :     : : :     : : :     : : :	
Db	1326	ESOLODTQELLBPETRKINLSRSIRQLSEBK-----NSLEQOESEEEBARNLEKQV	1378
		: : :     : : :     : : :     : : :     : : :     : : :	



Db 1548 ALEQVVEMRTOLEBELEDELOATEDAKLRLEVNMQAKMAQFERDLQTRDEQNEEKRLLI 1607

Qy 469 SEIDL--KLENSISLQKAAKAG-----NAEDVOHOILATESNQEVYRMILLDTQKSA 521

Db 1608 KQVRELEAELEDERKQALAVASKKWEIDLDKLEAQIEAANKARDEVIKRLKLAQMK 1667

Qy 522 LKTEIKETIVS----FLQ-----KITDQLNQLKQOE-----EDFRKQLED 558

Db 1668 DYQRELEEARSRDEIFAQSEKSEKSLKLEAEILQLQOEELASSERARRHAEQERDELAD 1727

Qy 559 E-----EGRKA---EK---ENTTAELTEINKWRLLVEELYNKTPQLQLDAREVEKQA 607

Db 1728 EITNSAKSALLDEARLEARIQAELEEEQSNMELLNDRPKTTQLQVDTLNAE--- 1784

Qy 608 LLNEHGAQOLNIRDSYAKLLGHON--LKOKIKHV-----VKLDKNSOLKSEVSKLR 660

Db 1785 LAERSAAQS-----DNARQLERQNKELKAKLOELGAVKSKFKATISALEAKIQLE 1839

Qy 661 CQL---AKKQSEYKLOEELNKVL-----GIKHFDPSKAPHHESKENF-ALKTPLK 707

Db 1840 EQLEQAEKERAANKLVRRTEKLEKTFMQVEDERRHADQYEQMEKANARMKQLKQLE 1899

Qy 708 EGNINCYRA 716

Db 1900 EABEATRA 1908

RESULT 15

AAU79590

ID AAU79590 standard; Protein; 1388 AA.

XX AC AAU79590;

XX DT 24-SEP-2002 (first entry)

XX DE Human kinesin motor protein, Hskif15.

XX KW Human; enzyme; cytosolic; neuroprotective; kinesin motor protein; Hskif15; microtubule stimulated ATPase activity; Hskif15M1A367;

KW Hskif15M1A367; kinesin; mitotic spindle; microtubule motor; antigen; therapeutic; diagnosis; cancer; neurological disorder; vesicular transport; atherosclerosis; tumour; abnormal wound healing; inflammatory disorder; immune disorder; rheumatoid arthritis; ocular angiogenic disease; glaucoma; cardiovascular disease; hypertension; diastolic dysfunction; fungal disease; aspergillosis.

XX OS Homo sapiens.

XX PN US6391613-B1.

XX PD 21-MAY-2002.

XX PF 27-NOV-2000; 2000US-0723219.

XX PR 04-JUN-1999; 99US-137423P.

XX PR 17-MAY-2000; 2000US-0572191.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C, Sakowicz R, Wood KW;

XX DR WPI; 2002-498776/53.

XX DR N-PSDB; ABK86978.

XX PT New nucleic acid encoding a human kinesin motor protein designated Hskif15 which has microtubule stimulated ATPase activity, for diagnosing and treating cancer, neurological disorders and disorders of vesicular transport -

XX PS Claim 1; Fig 2; 21pp; English.

XX CC The invention discloses an isolated nucleic acid encoding a human kinesin motor protein, Hskif15, which has microtubule stimulated ATPase activity.

CC and two truncated versions, Hskif15M1A367 and Hskif15M1A401, retaining the amino terminus. Hskif15, a member of the kinesin superfamily, has been found to be essential for mitotic spindle formation. The predicted structure comprises an amino-terminal kinesin-like microtubule "motor" domain. The polynucleotides and polypeptides may be used therapeutically to diagnose and prevent or treat cancer, neurological disorders and disorders of vesicular transport. Examples of the conditions that can be treated include atherosclerosis, tumours, abnormal wound healing, inflammatory and immune disorders (such as rheumatoid arthritis), ocular angiogenic disease (such as glaucoma), cardiovascular disease (such as hypertension), diastolic dysfunction and fungal disease (such as aspergillosis). The polynucleotides and polypeptides may also be used to screen for modulators of Hskif15 and raise antibodies. The sequence presented is the human kinesin motor protein, Hskif15.

XX Sequence 1388 AA;

Qy 37 SFKRSQRFKQ-----QKESQNLNVD-----K 58

Db 598 SKQYEFEFKELTRKQLESELOSLQKANKLNLENLEATKACKRQEVSQLNKHAEITLK 657

Qy 59 DTTLPASARKVKSSESKEKES-----QKNDKDLKILEK 90

Db 658 IITPTKAYQLHSRPVKLSPMGSGSLYTONSILNDILNEPVPPEMNEQAEASE 717

Qy 91 EIRVLLQERGAODRRIQD-----LETELEKME-----ARLNAAERKETSLS 131

Db 718 ELRTVQEQMSALQAKLDEEHKLNKLOQHVDKLHHSTQMOELFSSERIDWTQKQEBLLS 777

Qy 132 ANNATLEKQLIELTRTNELLKSPSE-----NGKNKLRILSLLELMKLRNKRRTKMRGM 186

Db 778 QLN-VLEKQLQETQTKNDFLKSEVHDLRVLHSAADKLSVSKVSEPKTQOERFENKL- 835

Qy 187 AKQEGMEKMLQVTKRSLE-----ESQ-----KTAQEGKLVSIIEKEKD 226

Db 836 -SERHMHVQLQDLNLRLENKLESKACLDQSYDNLQIMKEFIDQLSRNLFONFKENET 894

Qy 227 EKSTETKILLEYIEEISCASQDVKEYKLDIAOLEENKLEKNDEILSKLOSLEENIVILSKQ 286

Db 895 LKSDLNLMLELLE-----AEKERNKLSL-QFEEDKNSKEILLKVLAVRO-----EKQ 943

Qy 287 VEDLNVRKQ-----LLEKEKE-DHVNRRNHNENLNRAEMQNKLQKFILEQOE 333

Db 944 KE--TAKCEQQMAKVQKLEESLATEKRVISSLEKSRDSDDKKVVDLMNQIOEL-----R 995

Qy 334 EKLOQKELQIDSLQOEKELSSSLHOKLCSFQOEMVK----- 370

Db 996 SSVCEKTETIDTLKQELADINKYNSALVDREESVLLKQOEVDILDLETLRILSED 1055

Qy 371 -EKNLFEELKOTLDLQKLOQKQAEARLYKQLEEEAKSRAEELKLEELKLGAELE 429

Db 1056 IERDMLCEDLAHATEQLNMLTEASKHSGLLQSAQEELTKKEALIQELQHLNOKKEVE 1115

Qy 430 KSSAAHTQATILLQEKYDSMVQSLED--VTAQFESYKA---LTASEIEDLKLENSLQE 483

Db 1116 QKNEYNFKMRQLEHVMDSAAEDPQSPKTPPHFQTHLAKLLETQEQEIEDGRASKTSLH 1175

Qy 484 KAAKAGNAEDVQHOILATESNQEVYRMILLDQ---TKSALKETEIKEITVSFQKLTID 540

Db 1176 LVTKLNEDREVKNAILRMKEQLREMNLRLESQOLIEKNLLQOQLDDIK----RQKENS 1232

Qy 541 LQN-----QLKQOEEDFRKQ-----LEDEEGRKAERKENTFAETLEIN 578

Db 1233 DQNPDPNQLKNEQESIKERLAKSKIIVEEMLMKADLEEVSQALYNKMECLRMTDEVE 1292

Qy 579 KWRLLYEELYNKTRPFQQLDAFEVQKALLNEHGAQOEQLNKRDSYAKLLGHONLKQK 638

Db 1293 RTQTLSEKAFQEKQLRSLKEEMVEERERTSQEEMMLRQVECLAENGKLVGHQNLHOK 1352

Search completed: December 20, 2002, 15:11:17  
Job time : 45.4263 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:06:53 ; Search time 9.30678 Seconds  
(without alignments)  
2812.095 Million cell updates/sec

Title: US-09-685-010-48

Perfect score: 3104

Sequence: 1 MRALSLEMLKRNKRETKMR.....FCHASKENFTPLKEGNPNC 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	3104	100.0	794	1	HMNR_MOUSE
2	2091.5	67.4	498	1	HMNR_RAT
3	1949	62.8	724	1	HMNR_HUMAN
4	431.5	13.9	1790	1	USO1_YEAST
5	421.5	13.6	1939	1	MYH4_HUMAN
6	421	13.6	1938	1	MYH4_RABIT
7	417.5	13.5	1941	1	MYH2_HUMAN
8	415	13.4	1939	1	MYH1_HUMAN
9	413	13.4	1940	1	MYH3_RAT
10	413	13.3	1938	1	MYSS_CHICK
11	408	13.1	1130	1	YL17_CAEEL
12	406	13.1	1935	1	MYSS_CYPCA
13	404	13.0	1940	1	MYH3_CHICK
14	400.5	12.9	1084	1	MYSS_RABIT
15	400	12.9	1940	1	MYH3_HUMAN
16	396	12.8	1938	1	MYHD_HUMAN
17	393	12.7	1939	1	MYH6_MESAU
18	389.5	12.5	1935	1	MYH7_HUMAN
19	386.5	12.5	1939	1	MYH6_HUMAN
20	387	12.5	1938	1	MYH6_MOUSE
21	387	12.5	1938	1	MYH6_RAT
22	384	12.4	1960	1	MYH9_HUMAN
23	383.5	12.4	1976	1	MYHA_RAT
24	383	12.3	1427	1	REST_HUMAN
25	382.5	12.3	1102	1	MYSC_CHICK
26	379.5	12.2	1937	1	MYH8_HUMAN
27	379.5	12.2	1938	1	MYA_AEQIR
28	378.5	12.2	1934	1	MYH7_MESAU
29	377	12.1	1935	1	MYH7_RAT
30	376.5	12.1	1978	1	RA50_AQUAE
31	375.5	12.1	1976	1	MYHA_HUMAN
32	375	12.1	2230	1	GOG4_HUMAN
33	373.5	12.0	1961	1	MYH9_RAT

34	373.5	12.0	1976	1	MYHA_BOVIN
35	373	12.0	1978	1	MYHB_CHICK
36	371.5	12.0	1935	1	MYH7_PIG
37	366	11.8	1972	1	MYHB_RABIT
38	365	11.8	1433	1	REST_CHICK
39	364.5	11.7	2017	1	MYSN_DROME
40	363.5	11.7	944	1	NUF1_YEAST
41	363.5	11.7	2245	1	MYSJ_DICDI
42	363	11.7	1875	1	MLP1_YEAST
43	363	11.7	1972	1	MYHB_HUMAN
44	362.5	11.7	1966	1	MYSB_CAEEL
45	362	11.7	976	1	SCPI_HUMAN

Q27991 bos taurus  
P10587 gallus gall  
P79293 sus scrofa  
P35748 oryctolagus  
O42184 gallus gall  
Q99323 drosophila  
P32380 saccharomyc  
P54697 dictyosteli  
O02455 saccharomyc  
P35749 homo sapien  
P02566 caenorhabdi  
Q15431 homo sapien

ALIGNMENTS

RESULT 1

ID	HMNR_MOUSE	STANDARD;	PRT;	794 AA.
AC	O00547;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Hyaluronan mediated motility receptor (Intracellular hyaluronidic acid binding protein) (Receptor for hyaluronan-mediated motility).			
DE	HMNR OR IHABP OR RHAMM.			
GN	Mus musculus (Mouse).			
OS	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
[1]				
RN	SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.			
RC	TISSUE=Lung;			
RX	MEDLINE=98264863; PubMed=9601097;			
RA	Hofmann M., Fieber C., Assmann V., Goettlicher M., Sleeman J.,			
RA	Plug R., Howells N., von Stein O., Ponta H., Herrlich P.;			
RT	"Identification of IHABP, a 95 kDa intracellular hyaluronate binding protein."			
RT	J. Cell Sci. 111:1673-1684(1998).			
[2]				
RN	SEQUENCE FROM N.A.			
RP	Zhao Y., Zhang S., Turley E.;			
RA	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
[3]				
RN	SEQUENCE OF 1-183 FROM N.A., AND TISSUE SPECIFICITY.			
RP	STRAIN=129/Sv;			
RC	MEDLINE=99107769; PubMed=9889313;			
RA	Fieber C., Plug R., Sleeman J., Dall P., Ponta H., Hofmann M.;			
RT	"Characterization of the murine gene encoding the intracellular hyaluronan receptor IHABP."			
-RL	Gene 226:41-50(1999).			
[4]				
RN	SEQUENCE OF 164-794 FROM N.A., AND ALTERNATIVE SPLICING.			
RP	STRAIN=BALB/c; TISSUE=Fibroblast;			
RC	MEDLINE=96011639; PubMed=7590272;			
RX	Entwistle J., Zhang S., Yang B., Wong C., Li Q., Hall C.L., A J.,			
RA	Mowat M., Greenberg A.H., Turley E.A.;			
RT	"Characterization of the murine gene encoding the hyaluronan receptor RHAMM."			
RL	Gene 163:233-238(1995).			
[5]				
RN	SEQUENCE OF 318-794 FROM N.A., AND CHARACTERIZATION.			
RP	STRAIN=BALB/c;			
RC	MEDLINE=92299690; PubMed=1376732;			
RX	Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,			
RA	Cripps V., Austen L., Nance D.M., Turley E.A.;			
RT	"Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motility."			
-RL	J. Cell Biol. 117:1343-1350(1992).			
[6]				
RN	CHARACTERIZATION.			
RP	MEDLINE=94308286; PubMed=7518470;			

RA Hall C.L., Wang C., Lange L.A., Turley E.A.;  
 RT "Hyaluronan and the hyaluronan receptor RHAMM promote focal adhesion  
 RL turnover and transient tyrosine kinase activity.";  
 RN J. Cell Biol. 126:575-588(1994).  
 RP [7]  
 RX ERK REGULATION, AND SUBCELLULAR LOCATION.  
 RA MEDLINE=98225222; PubMed=9556628;  
 RA Zhang S., Chang M.C., Zylka D., Turley S., Harrison R., Turley E.A.;  
 RT "The hyaluronan receptor RHAMM regulates extracellular-regulated  
 RL kinase.";  
 RN J. Biol. Chem. 273:11342-11348(1998).  
 RP [8]  
 RX REVIEW.  
 RA MEDLINE=99059494; PubMed=9845361;  
 RA Hofmann M., Assmann V., Fieber C., Sleeman J.P., Moll J., Ponta H.,  
 RA Hart I.R., Herrlich P.;  
 RT "Problems with RHAMM: a new link between surface adhesion and  
 RT oncogenesis?";  
 RL Cell 95:591-592(1998).  
 CC -1- FUNCTION: INVOLVED IN CELL MOTILITY. WHEN HYALURONAN BINDS TO  
 CC HMMR, THE PHOSPHORYLATION OF A NUMBER OF PROTEINS, INCLUDING THE  
 CC FOCAL ADHESION KINASE OCCURS. MAY ALSO BE INVOLVED IN CELLULAR  
 CC TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING  
 CC EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY.  
 CC -1- SUBUNIT: SUBUNIT OF THE HARC COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: RHAMM1V4 (SHOWN HERE) AND  
 CC RHAMM1; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF031932; AAC12655.1; -;  
 DR EMBL; AF079222; AAD08670.1; -;  
 DR EMBL; AJ005919; CAA06768.1; -;  
 DR EMBL; AJ005920; CAA06768.1; JOINED.  
 DR EMBL; AJ005921; CAA06768.1; JOINED.  
 DR EMBL; AJ005922; CAA06768.1; JOINED.  
 DR EMBL; AJ005923; CAA06768.1; JOINED.  
 DR EMBL; AJ005924; CAA06768.1; JOINED.  
 DR EMBL; X64550; CAA45849.1; -;  
 DR EMBL; X64550; CAA45848.1; -;  
 DR MGD; MGI:104667; Hmrr.  
 KW Hyaluronic acid; Alternative splicing; Repeat; Glycoprotein.  
 FT DOMAIN 442 546 5 X 21 AA NEAR PERFECT TANDEM REPEATS.  
 FT REPEAT 442 462 1.  
 FT REPEAT 462 483 2.  
 FT REPEAT 483 504 3.  
 FT REPEAT 504 525 4.  
 FT REPEAT 525 546 5.  
 FT REPEAT 546 561 6.  
 FT DOMAIN 719 729  
 FT DOMAIN 719 750  
 FT CARBOHYD 53  
 FT CARBOHYD 134  
 FT CARBOHYD 279  
 FT CARBOHYD 446  
 FT CARBOHYD 467  
 FT CARBOHYD 488  
 FT CARBOHYD 509  
 FT CARBOHYD 530  
 FT CARBOHYD 561  
 FT CARBOHYD 601  
 FT VARSPLIC 218  
 FT CONFLICT 19  
 FT CONFLICT 55  
 FT CONFLICT 71  
 FT CONFLICT 89  
 FT CONFLICT 91  
 HYALURONIC ACID-BINDING (POTENTIAL).  
 HYALURONIC ACID-BINDING (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 MISSING (IN ISOFORM RHAMM1).  
 P -> Q (IN REF. 3).  
 S -> N (IN REF. 2).  
 K -> T (IN REF. 2).  
 EKE -> QKH (IN REF. 2).

FT CONFLICT 94 94 A -> V (IN REF. 3).  
 FT CONFLICT 540 540 S -> T (IN REF. 2).  
 FT CONFLICT 668 668 E -> D (IN REF. 2).  
 SQ SEQUENCE 794 AA; 91799 MW; 74DB3D236224499C CRC64;  
 Query Match 100.0%; Score 3104; DB 1; Length 794;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-98;  
 Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKQATQKDLTSEKGVILEGKLSIEKE 60  
 DB 164 MRALSLEMLKLNKRETKMRSMVMVQEGMELKQATQKDLTSEKGVILEGKLSIEKE 223  
 QY 61 KIDECETELLEYIQEISCASQDQVEKCKVDIAQLEEDLKEKDREILSLKQLEENITFS 120  
 DB 224 KIDECETELLEYIQEISCASQDQVEKCKVDIAQLEEDLKEKDREILSLKQLEENITFS 283  
 QY 121 KOIEDLTVKQQLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKLOQELQSQ 180  
 DB 284 KOIEDLTVKQQLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKLOQELQSQ 343  
 QY 181 SLQOEKELSLRLOQLCSFQSEMTSEKRVKEELKLAELDAVQOEQSERLVKQLE 240  
 DB 344 SLQOEKELSLRLOQLCSFQSEMTSEKRVKEELKLAELDAVQOEQSERLVKQLE 403  
 QY 241 EERKSTAEQLTRDLNLLREKEVELEKHTAAHAQAIIAQEKYNDTAQSLRDVTAQLE 300  
 DB 404 EERKSTAEQLTRDLNLLREKEVELEKHTAAHAQAIIAQEKYNDTAQSLRDVTAQLE 463  
 QY 301 EKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE 360  
 DB 464 EKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE 523  
 QY 361 SVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE 420  
 DB 524 SVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLE 583  
 QY 421 TAESTNQEYARVQDLQNRSTLKEEKEITSSFLKEITDKNQLRQDEDFRKLQLEEGK 480  
 DB 584 TAESTNQEYARVQDLQNRSTLKEEKEITSSFLKEITDKNQLRQDEDFRKLQLEEGK 643  
 QY 481 KRTAEKENVMTELTWEINKWRLLYELYEKTPFQOQQLDAFAEAKOALLNEHGATQEQLN 540  
 DB 644 KRTAEKENVMTELTWEINKWRLLYELYEKTPFQOQQLDAFAEAKOALLNEHGATQEQLN 703  
 QY 541 KIRDSYAQLLGHQNLKQIKHVVKLDKNSQLKSEVSKLRSOLVKRKONELRQLELQDKA 600  
 DB 704 KIRDSYAQLLGHQNLKQIKHVVKLDKNSQLKSEVSKLRSOLVKRKONELRQLELQDKA 763  
 QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631  
 DB 764 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 794  
 RESULT 2  
 HMMR\_RAT  
 ID HMMR\_RAT STANDARD; PRT; 498 AA.  
 AC P97779;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Hyaluronan mediated motility receptor (intracellular hyaluronan acid  
 DE binding protein) (Receptor for hyaluronan-mediated motility).  
 GN HMMR OR IHABP OR RHAMM  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID:10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Vascular smooth muscle;  
 RA Savani R.C., Hou G.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INVOLVED IN CELL MOTILITY. WHEN HYALURONAN BINDS TO  
CC HMMR, THE PHOSPHORYLATION OF A NUMBER OF PROTEINS, INCLUDING THE  
CC FOCAL ADHESION KINASE OCCURS. MAY ALSO BE INVOLVED IN CELLULAR  
CC TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING  
CC EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY (BY SIMILARITY).  
CC -!- SUBUNIT: SUBUNIT OF THE HARC COMPLEX.  
CC -!- SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC (BY  
CC SIMILARITY).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U87983; AAB47997.1; -;  
KW Hyaluronic acid; Repeat; Glycoprotein.  
FT DOMAIN 420 430 HYALURONIC ACID-BINDING (POTENTIAL).  
FT DOMAIN 442 451 HYALURONIC ACID-BINDING (POTENTIAL).  
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 498 AA; 57858 MW; 58037C79BD5C2A70 CRC64;  
  
Query Match 67.4%; Score 2091.5; DB 1; Length 498;  
Best Local Similarity 75.7%; Pred. No. 3.1e-64;  
Matches 432; Conservative 26; Mismatches 26; Indels 87; Gaps 2;  
  
Qy 64 EKCETKLEYIQBISCASDQVCKVDIAQLEEDLKEKDEILSKLOSENITFSKQI 123  
Db 12 EKSETEKLEVEIEBISCASDQVCKVDIAQLEEDLKEKDEILSKLOSEKVSFSKQI 71  
Qy 124 EDLVKCOLLTERDNLVSKDRRAETLSAEMQILTERALERQREYKLOQKELQSQSL 183  
Db 72 EDLVKCOLLEARDLVSKDRRAEUSLSAEMQVLTETKLLERQREYKLOQKELQSQSL 131  
Qy 184 QOEKLSARLOQQLCSFOEEMTSKKNVFKELKALAEALDAVQKEQSERLVKQLEER 243  
Db 132 QOEKLSARLOQQLCSFOEEMTSKKNVFKELKALAEALDAVQKEQSERLVKQLEER 191  
Qy 244 KSTAEQLRLDNLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQEKY 303  
Db 192 KSTAEQLRLDNLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQEKY 241  
Qy 304 NDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESVQ 363  
Db 242 ----- 241  
Qy 364 EKYNDTAQSLRDVSAQLESYKSSITLKETEDLKLNLTIQEKVMAEKSVEDVQOQILTA 423  
Db 242 -----AQLESYKSSITLKEEDLKLNLTIQEKVMAEKSVEDVQOQILTA 287  
Qy 424 STNOEYARMVDLQNRSTLKEEIKETTSFLEKITDLKQLRQOQEDFRKQLEEKGRKT 483  
Db 288 STNOEYARMVDLQNRSTLKEEIKETTSFLEKITDLKQLRQOQEDFRKQLEEKGRKT 347  
Qy 484 AEKENVTELTWENKRWLLYELYEYKTPQOQDLDAFAEKQALLNPHGATQOLNKIR 543  
Db 348 TEKETAFTVTELTWENKRWLLYELYEYKTPQOQDLDAFAEKQALLNPHGATQOLNKIR 407  
Qy 544 DSYAQLLGHQNLKOKIKHVKLKDENSQKSEVSKLSRQKQNEURLOGELDKALGI 603  
Db 408 DSYAQLLGHQNLKOKIKHVKLKDENSQKSEVSKLSRQKQNEURLOGELDKALGI 467  
Qy 604 RHDFPSKAFCHASKENF-----TPLKEGPNCC 631  
Db 468 RHDFPSKAFCHASKENVTLKTPLEKGNPCC 498

RESULT 3  
HMMR\_HUMAN

ID AC 075330; Q92767; STANDARD; PRT; 724 AA.  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hyaluronan mediated motility receptor (intracellular hyaluronic acid  
DE binding protein) (Receptor for hyaluronan-mediated motility) (CD168  
DE antigen).  
GN HMMR OR IHABP OR RHAMM.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.  
RC TISSUE-Breast carcinoma; PubMed=9601098;  
RA MEDLINE=96264864; PubMed=9601098;  
RT "the human hyaluronan receptor RHAMM is expressed as an intracellular  
RL protein in breast cancer cells."; J. Cell Sci. 111:1685-1694(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Breast;  
RA MEDLINE=97045829; PubMed=8890751;  
RT Wang C., Entwistle J., Hou G., Li Q., Turley E.A.;  
RT "The characterization of a human RHAMM CDNA: conservation of the  
RL hyaluronan-binding domains."; Gene 174:299-306(1996).  
CC -!- FUNCTION: INVOLVED IN CELL MOTILITY. WHEN HYALURONAN BINDS TO  
CC HMMR, THE PHOSPHORYLATION OF A NUMBER OF PROTEINS, INCLUDING THE  
CC FOCAL ADHESION KINASE OCCURS. MAY ALSO BE INVOLVED IN CELLULAR  
CC TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING  
CC EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY.  
CC -!- SUBUNIT: SUBUNIT OF THE HARC COMPLEX.  
CC -!- SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC (BY  
CC SIMILARITY).  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BREAST CANCER CELL LINES AND IN  
CC NORMAL BREAST TISSUE.  
CC -!- DATABASE: NAME=PROW; NOTE=PROW 2:76-84(2001);  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/802868666\_g.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF032862; AAC32548.1; -;  
DR EMBL; U29343; AAC52049.1; -;  
DR Genew; HGNC:5012; HMMR.  
DR MIN; 600936; -;  
KW Hyaluronic acid; Alternative splicing; Repeat; Glycoprotein; Antigen.  
FT DOMAIN 635 645 HYALURONIC ACID-BINDING (POTENTIAL).  
FT DOMAIN 657 666 HYALURONIC ACID-BINDING (POTENTIAL).  
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 588 588 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 75 90 MISSING (IN ISOFORM B).  
FT CONFLICT 75 75 K -> KK (IN REF. 2).  
FT CONFLICT 103 103 S -> R (IN REF. 2).  
FT CONFLICT 277 277 E -> D (IN REF. 2).  
FT CONFLICT 298 298 K -> T (IN REF. 2).  
FT CONFLICT 322 322 K -> E (IN REF. 2).  
FT CONFLICT 330 332 QER -> REH (IN REF. 2).  
SQ SEQUENCE 724 AA; 84031 MW; EA68AD6D2A626926 CRC64;  
  
Query Match 62.8%; Score 1949; DB 1; Length 724;

Best Local Similarity 64.8%; Pred. No. 2.6e-59;									
Matches 411; Conservative 48; Mismatches 87; Indels 88; Gaps 3;									
Qy	1	MRALSLEMLKRNKRTKMRSMVMKQEGMELKLOATQKDLTSGKIVOLEGKLVSIKKE	60						
Db	163	LRLSLLEMLKRNKRTKMRSMVMKQEGMELKLOATQKDLTSGKIVOLEGKLVSIKKE	222						
Qy	61	KIDKCETEKLEIYQEIASCADQVEKCKVDIAQLEEDLKEKDRILSLKQSLSEENIT-F	119						
Db	223	KIDKCETEKLEIYEEISCADQVEKCKVDIAQLEEDLKEKDRILSLKQSLSEENIT-L	282						
Qy	120	SKQIEDLTWKQLLETERNLVSKDRERAEITLSAEMQILTERKLAREOYEYKLOQKEIQS	179						
Db	283	SKQVEDLVKCOLLEKEKEDHVRNREHNENLNAEMQNLKQFLEQREKLOQKEIQI	342						
Qy	180	OSILOQEKELSLARLOOQLCSFOEEMTSEKNVFEELKLALAEALDAVQOKEESERLVKOL	239						
Db	343	DSLLQOKEKELSSSLQKLCFSFOEMVMEKNLEFEELKOTLDELKLOQKEEQEARLVKOL	402						
Qy	240	EEERKSTABQLTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESV	299						
Db	403	EEEAASRAEELKLEKLGKAELEKSSAHTQATLLQEKYDSMVQSLSDVT-----	456						
Qy	300	QEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQ	359						
Db	457	-----	456						
Qy	360	ESVQEKYNDTAQSLRDVSAQLESYKSSITLKEITLDELKLENITLQEKVMAEKSVEDVQOOI	419						
Db	457	-----AQFESYKALTASEIEDLKLSENSQEKAAKAGNAEDVQHOI	498						
Qy	420	LTAESTNOEYARMVDLQNRSLKBEETKEITSSPLEKITDLKNLROQDEDFRKLQBEK	479						
Db	499	LATESNQEVYRMLDLQTPKSALKETEITVSLQKITDLQNLKQBEEDFRKLQLEDE	558						
Qy	480	GKRTAEKENVWTEIMINKWRLLYEELYEYKYPFQQQLDAFEAEKQALLNHHGATQSQL	539						
Db	559	EGRKAKEKENTTAELEINKNRLLYEELYEYKYPFQQQLDAFEAEKQALLNHHGAAQEQ	618						
Qy	540	NKIRSYAOLLGHONLKQIKHVVKLKDNLSOLKSEVSKRSOLVKRKONELRLQELQDK	599						
Db	619	NKIRSYAKLGHQNLKQIKHVVKLKDNLSOLKSEVSKRSOLVKRKONELRLQELQDK	678						
Qy	600	ALGIRHFDPSKAFCHASKENF---TPLKEGNPNC	630						
Db	679	VLGIKHFDPKAFHSHESKENFALKTPKRGNTNC	712						
RESULT 4									
USOL_YEAST	STANDARD; PRT: 1790 AA.								
AC	P25386;								
DT	01-MAY-1992 (Rel. 22, Created)								
DT	01-MAY-1992 (Rel. 22, Last sequence update)								
DE	16-OCT-2001 (Rel. 40, Last annotation update)								
DE	Intracellular protein transport protein USOL.								
GN	USOL OR YDL058W.								
OS	Saccharomyces cerevisiae (Baker's yeast).								
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;								
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.								
OX	NCBI_TaxID=4932;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=X2180-1A;								
RX	MEDLINE=911185402; PubMed=2010462;								
RA	Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,								
RA	Yamasaki M.;								
RT	"A cytoskeleton-related gene, usol, is required for intracellular								
RT	protein transport in Saccharomyces cerevisiae.";								
RL	J. Cell Biol. 113:245-260(1991).								
RN	[2]								
RP	SEQUENCE OF 782-1790 FROM N.A.								
RA	Hostettler M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,								

Kendrick K.E.;	Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.	
[3]		
Bai Y., Symington L.S.;	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.	
-!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI COMPLEX.		
-!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE ER AND THE GOLGI COMPLEX.		
-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.		
-!- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.		
-----		
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
-----		
EMBL; X54378; CAA38253.1; -		
EMBL; L03188; AAB00143.1; -		
EMBL; U53668; AAB66659.1; -		
PIR; A38455; A38455.		
SGD; S0002216; USOL.		
InterPro; IP0002017; Spectrin.		
Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.		
DOMAIN 1 724		
DOMAIN 725 1790		
DOMAIN 465 487		
DOMAIN 991 1790		
DOMAIN 1172 1786		
CONFLICT 847 847		
CONFLICT 924 924		
CONFLICT 1253 1253		
CONFLICT 1319 1319		
CONFLICT 1461 1461		
CONFLICT 1581 1581		
CONFLICT 1600 1600		
CONFLICT 1661 1661		
CONFLICT 1772 1772		
CONFLICT 1790 AA; 206424 MW; 6CEB2B216E9FD4818 CRC64;		
SEQUENCE		
Query Match	13.9%; Score 431.5; DB 1; Length 1790;	
Best Local Similarity	24.2%; Pred. No. 4.1e-08;	
Matches 179; Conservative 149; Mismatches 236; Indels 177; Gaps 33;		
Qy	10	KLNRKRETKMRSMVKGEMELKLOATQKDLTSGKIVOLEGKLVSTEKEIDKCE
Db	798	ELKNVRDS-LDEMTQLRDVLETKDKKENQTALLEYKSTHKOEDSITKLEK-----
Qy	70	KLLEYIQEISCASQVQKCKVDIAQLEEDL-----KEKD-----REILSLKQ
Db	847	GLETLSQKKKAEDGINKMGKDLFALSREMQAVEBCKNLQKDKSNVNHQKTKSLKE
Qy	112	SLEENIT----FSKOIEDLTWKQLLETER-----DNLVSKDRERAE
Db	907	DIAAKITEIKAINENLEEMKTCNNLSKEKEHISKELVEYKSRFQSHDNVAKLTKLS
Qy	151	LS---AEMQILTERL--ALERQVEE-KLQOKELOSO-SLQOKEKLSARLOOQLCSFOE
Db	967	LANNYKDMAQENESLIKAVEKSNESLIQSLNQLNKLIDMSQOKE-----NFQIE
Qy	204	MTS-EKNVFEELKLALAEALDAVQOQ-----EQSERLVKOLEEERKSTAEQL
Db	1017	RGSIEKNI--EQLKKTISDLEQTEKEEIIKSSDKDEYESQISLLKLEKLETATTAND
Qy	251	TRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQO-----KY



```
Db 1353 EBEAKAEQLQRMKANSEVAQWRTKYETDAIQTTELEAKKKLAQLQDAEEHVEAVN 1412
QY 520 ----AFAEAKQALLNE-----HGATQEQOLN 540
Db 1413 SKCASLEKTKQRLQNEVEDLMDVRSNACIALDKKORNFQKVLAEWKQKYETQAELE 1472
QY 541 -----KIRDSYAQLLGH-----ONLKQKI-----KHVVKLKD 567
Db 1473 ASQKESRSLSTELFKVNAYBESLDHLETKRENKLNQAEISDLTQEAEGGKHITHELEK 1532
QY 568 ENSQLKSEVSKLRSOLVKKQNELRLQGLDKALGIR 604
Db 1533 VKQQLDHEKSELOSL---EEAEASLEHEEGKILRIQ 1566

RESULT 6
MYH4_RABIT
ID MYH4_RABIT STANDARD; PRT: 1938 AA.
AC Q28641;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Myosin heavy chain, skeletal muscle, juvenile.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Skeletal muscle;
RA Maeda K., Hostinova E., Roesch-Kleinhauf A., Schuster H., Gasperik J.,
RA Wittinghofer A.;
RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit
RT skeletal muscle and a novel cosynthesis of S-1 fragment with the
RT essential and regulatory light chains.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HWM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32574; AAA74199.1; -.
CC DR HSP3; P13538; 2MYS.
CC DR InterPro: IPR000048; IQ_region.
CC DR InterPro: IPR004009; Myosin_N.
CC DR InterPro: IPR002928; Myosin_tail.
CC DR InterPro: IPR001609; myosin_head.
CC DR Pfam: PF00063; myosin_head; 1.
CC DR Pfam: PF00612; IQ; 2.
CC DR Pfam: PF01576; Myosin_tail; 1.
CC DR Pfam: PF02736; Myosin_N; 1.
CC DR PRINTS: PR00193; MYOSINHEAVY.
CC DR ProDom: PD000355; myosin_head; 1.
CC DR SMART: SM00015; IQ; 1.
CC DR SMART: SM00242; MYSC; 1.
```

```
DR PROSITE; PS0096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 783 MYOSIN HEAD-LIKE.
FT DOMAIN 784 813 IQ.
FT DOMAIN 842 1938 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 658 680 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 760 774 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
FT MOD_RES 130 130 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD_RES 552 552 METHYLATION (SH-1) (BY SIMILARITY).
FT MOD_RES 756 756 METHYLATION (SH-1) (BY SIMILARITY).
FT MOD_RES 698 698 ALKYLATION (SH-2) (BY SIMILARITY).
FT MOD_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).
SQ SEQUENCE 1938 AA; D8A8A2EC5B182626 CRC64;

Query Match 13.6%; Score 421; DB 1; Length 1938;
Best Local Similarity 23.0%; Pred. No. 9.9e-08;
Matches 170; Conservative 150; Mismatches 251; Indels 168; Gaps 25;

QY 9 MKLNRKRETKMRSMVQOE--GMELKLAQTOKDLTESGKIVQLGKLVSTIEKEKIDEKC 66
DB 833 MKLYFKIKPLLSKSAETKEMANMKDEFKTESLAKAEKAEKELEKVMALQKNDLQL 892
QY 67 ETEKLEVIQIISCASDOVEKCV-----DIAOLEEDLKEKREILSLKQSLSEENIT- 118
DB 893 QVQAEADSLADAEPCDOLIKTKLEAKIKEVTERAEDEEEINAEELTAKRKLDECESE 952
QY 119 FSKQIEDTVKCOLLETERDNLVSKDRERATLSAEMOILTERLALERQEKYKQKELQ 178
DB 953 LKQDIDDELTLAKVEKHKATENK---VKNLTSEMAGLDETIAKTKKKKALQEAHQ 1008
QY 179 SQSLLOKE-----LSARLOOLCSFQEMTSEKNV-----FKEELKAL-A 220
DB 1009 TLDDLAEDKVNLTAKTKLEQVDDLEGSLEQEKIRMDLERAKRKKLGGDLKLAES 1069
QY 221 ELDAVQKQEEQSERLVQ-----LE---EERKSTAEQTLRLDNLREKEVELEKHAH 271
DB 1069 TMDIENDKOOLDKLLKKEFMSNLSQKIEDEQALAMQLKKIKELQARIELEEIEAE 1128
QY 272 QAAILIAQEKYNDPAQSLRDVTAQLESVQEKYNDPAQSLRDVTAQLESQEKYNDTAQSL 331
DB 1129 RASRAKAEKQKRSLSRELEISERLE-----EAGGATSAQIENMKKREAEFQKMR 1178
QY 332 RDVTAQLESQEKYNDPAQSLRDVTAQLESVQEKYNDPAQSLRDVTAQLESVKSTLKEI 391
DB 1179 RD----LEEAFLQHEATAATLURKHA--DSVAE-LGEQIDNLQVQKLEKESLKKMEI 1231
QY 392 EDLKLLENITLQEKVAMAEKSVEDVQQQLTAESTNQEYARMVQDQLQ-NRSTLKEE----- 445
DB 1232 DLLASNMTVSKAKGNLEKCMKRTLEDQVSELKTKREEHQRLINDLSAQARLARLQESGEFS 1291
QY 446 -----EIKETSSFLKKTDLKNQLRQ-----DED-FRKQLE 477
DB 1292 RQLDEKDSLVSQSRGKAQFTQQIEELKRLQEEETKAKSALAHALQSAHRHCDLLREQYE 1351
QY 478 EKGRTAKENVMTPELTMEINKRLLY-----EELYKTKPQQQLD----- 519
DB 1352 EEQAKAEQLQRMKANSEVAQWRTKYETDAIQTTELEAKKKLAQLQDAEEHVEAVN 1411
QY 520 ----AFAEAKQALLNE-----HGATQEQOLN-----KIRDSYAQLL 550
DB 1412 AKCASLEKTKQRLQNEVEDLMDVRSNACIALDKKORNFQKVLAEWKQKYETQAELE 1471
QY 551 GHQNLKQ-----KIKHVVK-----LKDENSQLKSEVSKLRSQ----- 583
DB 1472 ASQKESRSLSTELFKVNAYBESLDQELTKRENKLNQAEISDLTQEAEGGKHITHELEK 1531
QY 584 VKR--KONELRLQGLDKA 600
DB 1532 VKQVEQKSELQAALAEA 1550
```

DR	InterPro: IPR004009; Myosin_N.	
DR	InterPro: IPR002928; Myosin_tail.	
DR	InterPro: IPR001609; myosin_head.	
DR	Pfam: PF00063; myosin_head; 1.	
DR	Pfam: PF00612; IQ; 2.	
DR	Pfam: PF01576; Myosin_tail; 1.	
DR	Pfam: PF02736; Myosin_N; 1.	
DR	PRINTS: PR00193; MYOSINHEAVT.	
DR	PRODOM: PD000355; myosin_head; 1.	
DR	SMART: SM00015; IQ; 1.	
DR	SMART: SM00242; MYSC; 1.	
DR	PROSITE: PS00096; IQ; 1.	
KW	Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;	
KW	Calmodulin-binding; ATP-binding; Methylation; Alkylation.	
KW	Multigene family.	
FT	DOMAIN 1 786	MYOSIN HEAD-LIKE.
FT	DOMAIN 787 816	IQ.
FT	DOMAIN 845 1941	COILED COIL (POTENTIAL).
FT	NP_BIND 179 186	ATP (POTENTIAL).
FT	CONFLICT 1844 1844	K -> R (in ref. 3).
FT	SEQUENCE 1941 AA; 223043 MW; 681E866F83AE83F CRC64;	
QY	Query Match	13.5%; Score 417.5; DB 1; Length 1941;
QY	Best Local Similarity	22.6%; Pred. No. 1.3e-07;
QY	Matches 168; Conservative 145; Mismatches 258; Indels 171; Gaps	
QY	4	LSLEMLKRNKRETKMRSMVMVQEGMELKQATQKDLTESKGIIVQLSGKLVISIEKEID 63
DB	838	LFFKIRPLKSAETEKEMATKEEFQIK-----DELAKSEAKRKELEKMWTLKEND 892
QY	64	EKCETEKLLVEIQEISCASDOVEKCV-----DIAQLEEDLKEKDREIISLKSOSLEEN 116
DB	893	LQLQVQAAEAGLADAEERCIDLIKTIKLEAKIKEVTERAEDEEINAEITAKKRKLEDE 952
QY	117	IT-FSQIEDLTVKQLLETERDNLYSKDRERAETLSAEQMILTERLALEREQYEKQLQK 175
DB	953	CSELKDKIDDLLETLAKVKEKEHATENK-----VKNLTTEEMAGLDETIATLKTEKKALQEA 1008
QY	176	ELQSLSLQOKE-----LSARLQOQLCSFOEEMTSEKNV-----FKELKLA 218
DB	1009	HQOTLDLQAEEDKVNLTFKAKIKLSQVDDLEGSLEQEKLRMDLERAKKRKLEGDKLA 1068
QY	219	LAELDAVQOKEQSERLIVKOLE-----PERKSTAEQILRLDNLRLREKEVELEKHI 268
DB	1069	QESIMDIENEKQDLEKLLKKEPEISNLOSKTEDEQALGIQOKKIKELQARIELEEEI 1128
QY	269	AHAQAILLIAOEKYNDTAOSLRDVTQALESVOEKYNDTAQSLRDVTAQLESQEKYNDTA 328
DB	1129	EAFRASAKAEKORSLSRELEESIRLE-----EAGGATSAEIEMKKKEAEFQ 1178
QY	329	QSLRDVTAQLESQEKYNDTAOSLRDVTQALESVOEKYNDTAQSLRDVSAQLESYKSSLT 388
DB	1179	KMRD---LEEATLQHEATAATLRKKHA---DSVAE-LGEQIDNLRVQKLEKEKSEMK 1231
QY	389	KEIEDLKLENLTLOEKVAMAEKSVEDVQOQIILTAESTNOEYARMVQDL-QNRSTLKEE-- 445
DB	1232	MEIDDLASNVETYSKAKGNLEKMCRTLQDLSELKSEEEQQRLLNDLTAQGRQLQTESG 1291
QY	446	-----EIKETSSPLEKITDLKNQLRQO-----DD-PRK 474
DB	1292	EFSRQDKEALVYSQSRGQAQFTQIIEELKRLQLEEEIKAKNALAHALQSSRHDCDLRE 1351
QY	475	QLSEKGRRTAEKENVMTLMEINKRRLY-----EELYEKTKPQQQLDA----- 520
DB	1352	QYEEQESKAEQALQALUSKANTEVAQWRKYETDAQRTLEEELAAKKLAQRQAEEHVE 1411
QY	521	-----FEAEKQALLNE-----HGATQEQLN-----KIRDSA 547
DB	1412	AVNACASLEKTKQRQLONEVEDLMDVETRNAACALDKQRNFDKILAEWKQRCETHA 1471
QY	548	QL-----LGHQMLK-----QKIKHVVKLKDNSOLKSEVSKLSQSL----- 583
DB	1472	ELBASQEARSLTELFKIKNAYEESLDQLETLKRNKNLQOEISDLTEQIAEGGRKRIHE 1531



```

584  ---VKR---KONELRLQGLDKA 600
      :|:  |:  |:  |::|
Db 1532 LEKIKQVEQEKCELQAALSEA 1553

RESULT 8
MYHI_HUMAN  STANDARD; PRT; 1939 AA.
AC P12882; Q9Y622;
AT 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, adult 1 (Myosin heavy chain
DE Iix/d) (MyHC-Iix/d).
GN MYH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RC MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
RT heavy chain family: Implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RN SEQUENCE OF 1064-1939 FROM N.A.
RA MEDLINE=86176778; PubMed=2421254;
RA Saez L., Leinwand L.A.;
RT "Characterization of diverse forms of myosin heavy chain expressed in
RT adult human skeletal muscle.";
RL Nucleic Acids Res. 14:2951-2969(1986).
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF111785; AAD29951.1; -.
CC EMBL; X03740; CAA27380.1; -.
CC PIR; A23767; A23767.
CC HSP; P13538; ZWYS.
CC Genew; HGNC:7567; MYH1.
CC MIM; 160730; -.
CC InterPro; IPR000048; IQ_region.
CC InterPro; IPR004009; Myosin_N.
CC InterPro; IPR002928; Myosin_tail.
CC InterPro; IPR001609; myosin_head.
CC Pfam; PF00663; myosin_head; 1.
CC Pfam; PF00612; IQ; 2.
CC Pfam; PF01576; Myosin_tail; 1.

```



```

QY 443 KEE-----EIKEITSSFEKITDKNLQRO-----DE 470
Db 1285 QTEGEYSQRLDEKDTVLSGRGQAFQTQIEELKRLQEEBKAKSALAHALQSSRHDC 1344
QY 471 D-FRQLEBKGRFAEKNVTELTWENKWRLLY-----EELYEKTRPFQOOLD-- 519
Db 1345 DLLREQYEEQAKAELQRAMSKANSEVAQWRKYETDAIQRTETELEEAKKKLAQLQDA 1404
QY 520 -----AFFAEKQALINE-----HGATQROLN-----KI 542
Db 1405 EHVAVNAKASLEKTKQRLQNEVEDLMIDVERTNACAALDKKQRFDKILAEWKQK 1464
QY 543 RDSYAQLGHQNLKQ-----KIKHVVK-----LKDENSQKSEVSKLRSQ----- 583
Db 1465 EETHAELEASQKESRSLSTFLFKIKNAVEESLDQLETKRENKNIQEQEISDLTEQIAEGG 1524
QY 584 -----VKR--KQNELRLQGLDKA 600
Db 1525 KRIHELEKIKKQVEQEKSEKLAALAEA 1551

RESULT 9
MYH3_RAT
ID MYH3_RAT STANDARD: PRT; 1940 AA.
AC P12847;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87060988; PubMed-3783701;
RA Strehler E.E., Strehler-Page M.-A., Perriard J.C., Periasamy M.,
RA Nadal-Ginard B.
RT "Complete nucleotide and encoded amino acid sequence of a mammalian
RT myosin heavy chain gene. Evidence against intron-dependent evolution
RT of the rod."
RL J. Mol. Biol. 190:291-317(1986).
RC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X04267; CAA27817.1; -.
CC PIR: A24922; A24922.
CC HSSP: P13538; 2MY5.
CC InterPro: IPR000048; IQ_region.
CC InterPro: IPR004009; Myosin_N.

```

```

DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF01576; Myosin_tail; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IQ.
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 656 678 ACTIN-BINDING.
FT DOMAIN 758 772 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 696 696 ALKYLATION (SH-1).
FT MOD_RES 706 706 ALKYLATION (SH-2).
SQ SEQUENCE 1940 AA; 223857 MW; B5D546A596E5A696 CRC64;

Query Match 13.4%; Score 415; DB 1; Length 1940;
Best Local Similarity 22.0%; Pred. No. 1.6e-07;
Matches 165; Conservative 144; Mismatches 24; Indels 198; Gaps 22;

QY 4 LSLELMKLRNKRRETKMRSMVQEGMELKQATQKDLTSEKGIQVQLEKGLVSIKEKID 63
Db 833 LFFFKIKPLLSAETEKEMATKEE-----FQTKDELAKSEAKRKELEKLVTLQVEKD 887
QY 64 -----EKCETEKLLIYQIEISCASDVQVEKCV-----DIALEEDLKEKDEILSLKOS 112
Db 888 LQLOVQAENSELLD--AEERC--DOLIRAKFQLEAKIKVEKTEAEDEEINAEITAKKK 943
QY 113 LEENIT-FSKQLEDLTQKQLLETRDNLVSKDRERAETLSAEMOITLTERLALEQVEYEK 171
Db 944 LEDECELSKDKIDDLLELTAKVEKEKHATENK----VKNLTEELAGLDETIAKLTRKKA 999
QY 172 LQOKELQSLSLQBE-----LSARLQQLCSFQEEEMTSEKNV-----FKPE 214
Db 1000 LQEAHQITLDDLAQAEEDKVNLSKLSKLEQVDDLESSLEQEKLRVLDLERNKRLK 1059
QY 215 LKALAE-LDAVQVQKEQESERLVK-----QLE---EERKSTAEQLTRLDNLLREKEVEL 264
Db 1060 LKLAQESILDLNDKQDLERLKKDKDFEYSQLSQKVEDQETLSLQLOKKIKELQARIEEL 1119
QY 265 EKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQOE-K 323
Db 1120 EEEIEAERATRAKTRAKTRKQSDYARELEUESERLEE-----AGVTSTQIEUNKKRAE 1171
QY 324 YNDTAQSLRDVTAQLESE-----QEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQ 379
Db 1172 FLKLRDLEEAATLQHEATVATLRKKHADSAAELAQIDNLRVKQK-----1217
QY 380 LESYKSTLKELEDKLEMLTLQEKVMAEKSVEDVQOQLTAESTNQYARMVQDLQNR 439
Db 1218 LEKESEFKLEIDDLSSSVESVSKANLEKICRTLEDQLSSEARGKNEETQRSLSLTQ 1277
QY 440 STLKEEIKEIT-----SSFEKITDKNLQROODE-----471
Db 1278 KSRLOTEAGELSRQLEEKESIVSLSRQKQAFQOIEELKRLQEEENKAKNALAHALQSS 1337
QY 472 -----FRKQLEBKGRFAEKNVTELTWENKWRLLY-----EELYEKTRPFQO 517
Db 1338 RHDCDLLREQYEEQEGKAEQALQALSKANSEVAQWRKYETDAIQRTETELEEAKKKLAQR 1397
QY 518 L-----518
Db 1398 LQDSEEQVEAVNAKASLEKTKQRLQGEVEDLMVDVERANSALAAALDKKQRFDKVLAEW 1457

```

QY 519 ----DAFEAEKQALLNEHGATQQLNKIRDSYAQLGHONLKQKIKHVVKLRKDNSQLKS 574  
 Db 1458 KTKCEESQAELEAALKSRSLSTFLFKLNAY-----EEALDQLETVKRENKLNLEQ 1508  
 QY 575 EVSKLRSLVKRQN-----ELRLQGELEKA 600  
 Db 1509 ETADLTEQIAENGKSTHELEKSRKQMELEKA 1539

RESULT 10  
 MYSS\_CHICK  
 ID MYSS\_CHICK STANDARD; PRT: 1938 AA.  
 AC P13538; O13228;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, skeletal muscle, adult.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White leghorn; TISSUE=Pectoralis muscle;  
 RA Chao T.H., Bandman E., Moore L.;  
 RT "Cloning, nucleotide sequence and characterization of a full-length  
 RT cDNA encoding the myosin heavy chain from adult chicken pectoralis  
 RT major muscle.";  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DDAJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-205.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041767; PubMed=1939027;  
 RA Hayashida M., Maita T., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: I.  
 RT Sequence of the amino-terminal 23 kDa fragment.";  
 RL J. Biochem. 110:54-59(1991).  
 RN [3]  
 RP SEQUENCE OF 206-636.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041768; PubMed=1939028;  
 RA Komine Y., Maita T., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: II.  
 RT Sequence of the 50 kDa fragment of subfragment-1.";  
 RL J. Biochem. 110:60-67(1991).  
 RN [4]  
 RP SEQUENCE OF 637-837.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041769; PubMed=1939029;  
 RA Maita T., Miyaniishi T., Matsuzono K., Tanioka Y., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: III.  
 RT Sequence of the 22 kDa fragment and the alignment of the 23 kDa, 50  
 RT kDa, and 22 kDa fragments.";  
 RL J. Biochem. 110:68-74(1991).  
 RN [5]  
 RP SEQUENCE OF 838-1938.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041770; PubMed=1939030;  
 RA Maita T., Yajima E., Nagata S., Miyaniishi T., Nakayama S., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: IV.  
 RT Sequence of the rod, and the complete 1,938-residue sequence of the  
 RT heavy chain.";  
 RL J. Biochem. 110:75-87(1991).  
 RN [6]  
 RP PRELIMINARY SEQUENCE OF 1-808.  
 RX MEDLINE=87092420; PubMed=3467365;  
 RA Maita T., Hayashida M., Tanioka Y., Komine Y., Matsuda G.;  
 RT "The primary structure of the myosin head.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:416-420(1987).  
 RN [7]  
 RP SEQUENCE OF 842-1270.

RX MEDLINE=90121764; PubMed=2610940;  
 RA Watanabe B.;  
 RT "Complete amino-acid sequence of subfragment-2 in adult chicken  
 RT skeletal muscle myosin.";  
 RL Biol. Chem. Hoppe-Seyler 370:1027-1034(1989).  
 RN [8]  
 RP SEQUENCE OF 852-1108.  
 RX MEDLINE=89374803; PubMed=2775482;  
 RA Watanabe B.;  
 RT "Amino-acid sequence of the short subfragment-2 in adult chicken  
 RT skeletal muscle myosin.";  
 RL Biol. Chem. Hoppe-Seyler 370:549-558(1989).  
 RN [9]  
 RP SEQUENCE OF 1145-1270.  
 RX MEDLINE=89228549; PubMed=2713098;  
 RA Watanabe B.;  
 RT "Amino-acid sequence of the hinge region in chicken myosin  
 RT subfragment-2.";  
 RL Biol. Chem. Hoppe-Seyler 370:55-61(1989).  
 RN [10]  
 RP SEQUENCE OF 1857-1938 FROM N.A.  
 RX MEDLINE=87217964; PubMed=3034534;  
 RA Moriarity D.M., Barringer K.J., Dodgson J.B., Richter H.F.,  
 RA Young R.B.;  
 RT "Genomic clones encoding chicken myosin heavy-chain genes.";  
 RL DNA 6:91-99(1987).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-843.  
 RX MEDLINE=93303624; PubMed=8316857;  
 RA Rayment I., Rypniewski W.R., Schmidt-Base K., Smith R.,  
 RA Tomchick D.R., Benning M.M., Winkelmann D.A., Wesenberg G.,  
 RA Holden H.M.;  
 RT "Three-dimensional structure of myosin subfragment-1: a molecular  
 RT motor.";  
 RL Science 261:50-58(1993).  
 CC -!- FUNCTION: MUSCLE CONTRACTION. MYOSIN IS A PROTEIN THAT BINDS TO  
 CC F-ACTIN AND HAS ATPASE ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U87231; AAB47555.1; -;  
 DR EMBL; M16557; AAA48970.1; -;  
 DR PIR; PX0050; PX0051.  
 DR PIR; A26821; A26821.  
 DR PIR; S02082; S02082.  
 DR PIR; S04501; S04501.  
 DR PIR; S05515; S05515.  
 DR PDB; 2MYS; 11-JAN-97.  
 DR InterPro; IPR000048; IQ\_region.  
 DR InterPro; IPR004009; Myosin\_N.  
 DR InterPro; IPR002928; Myosin\_tail.  
 DR InterPro; IPR001609; myosin\_head.

DR Pfam: PF00063; myosin\_head; 1.  
DR Pfam: PF00612; IQ; 2.  
DR Pfam: PF01576; Myosin\_tail; 1.  
DR Pfam: PF02736; Myosin\_N; 1.  
DR ProDom: PD000355; myosin\_head; 1.  
DR SMART: SM00015; IQ; 1.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PS00096; IQ; 1.  
KW Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW ATP-binding; Methylation; Alkylation; Phosphorylation; Acetylation;  
KW Calmodulin-binding; Multigene family; 3D-structure.  
FT INIT\_MET 0 0  
FT DOMAIN 1 782 MYOSIN HEAD-LIKE.  
FT DOMAIN 783 812 IQ.  
FT DOMAIN 838 840 HINGE.  
FT DOMAIN 841 1938 COILED COIL (POTENTIAL).  
FT NP\_BIND 179 186 ATP (POTENTIAL).  
FT DOMAIN 657 679 ACTIN-BINDING.  
FT DOMAIN 759 773 ACTIN-BINDING.  
FT MOD\_RES 1 1 ACETYLATION.  
FT MOD\_RES 35 35 METHYLATION (MONO-).  
FT MOD\_RES 130 130 METHYLATION (TRI-).  
FT MOD\_RES 551 551 METHYLATION (TRI-).  
FT MOD\_RES 755 755 METHYLATION (MONO-).  
FT MOD\_RES 697 697 ALKYLATION (SH-1).  
FT MOD\_RES 707 707 ALKYLATION (SH-2).  
FT CONFLICT 907 907 C -> Q (IN REF. 7 AND 8).  
FT CONFLICT 980 980 L -> F (IN REF. 1).  
FT CONFLICT 1343 1343 E -> D (IN REF. 5).  
FT CONFLICT 1545 1545 S -> A (IN REF. 5).  
FT CONFLICT 1796 1797 HV -> QL (IN REF. 5).  
FT CONFLICT 1830 1830 S -> A (IN REF. 5).  
FT CONFLICT 1863 1863 I -> V (IN REF. 10).  
FT CONFLICT 1929 1931 IHG -> FH (IN REF. 10).  
SQ SEQUENCE 1938 AA; 223013 MW; EDD0ICEA2681E10F CRC64;

Query Match 13.3%; Score 413; DB 1; Length 1938;  
Best Local Similarity 22.6%; Pred. No. 1.8e-07;  
Matches 171; Conservative 133; Mismatches 249; Indels 204; Gaps 24;

QY 9 MGLRNKRTKMRSMVMKQF--GMELKLAQATQDLTSGKIVQLEGKLVSTKEKIDKDC 66  
DB 832 MKLFKIKPLKLSAESEKEMANMKFEFKTEELAKSAKEELKEMVLLQEKNDLQ 891  
QY 67 ETEKLEYIQTSCASDOVERKCV-----DIAQLEDLKEKDEILSLKQSPLENTT- 118  
DB 892 QVQAEADSLADAEKCDLIKTIOLEAKIKEVTERADEEIEINAEITAKRKLEDCSE 951  
QY 119 FSKQIEDLVKQCLETERDNLVSKDRERATLSAEMOILTERLALEREQYEKLOQKELQ 178  
DB 952 LKDDIDLELTAKVEKEKHATENK-----VKNLTTEMAVLDETIAKLTKEKALQEAHQ 1007  
QY 179 SOSLLQOEKE-----LSARLQOOLCSFOEMTSEKNV----- 210  
DB 1008 TLDDLQVEEDKVNLTAKTKLEQVDDLEGSLEQEKRLMDLERAKRKLEGLDLKLAHDS 1067  
QY 211 -----FKEELKALAEIDAVOQKEQSERLVKQLEERKSTABQTLRLDNLREK 260  
DB 1068 IMDLENDQQLDEKLKKDDFELSIOQSKIEDEQALGMQLQKKIK---ELQARIEL--EE 1122  
QY 261 EYELEKHHAAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESE 320  
DB 1123 ETEAERTSRKA-----EKHR-----ADLSRELEISERLEAGGA---TAAQIEMN 1166  
QY 321 QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLR-----DVTQALESVQEKYNDTAQSLRDV 376  
DB 1167 KKREAEFQKMRD-----LEEATLQHEATAAALRKHKHADSTAE-----GEQIDNLQRV 1215  
QY 377 SAQLESYSSTLTETEDLKENLTLOEKVANAQK---SVEDVQOQILTAESTNQEYARWV 433  
DB 1216 KQLEKESEKLMELDDLSNMSVSKAKANLEKRCMTFDQLSEIKTKESQNG---RMI 1272  
QY 434 QDLQNRSTLKKEEIKET-----SSPLEKITDLKNQL----- 465

DB 1273 NDLNTOARLQTTETGEYSRQAEKDALISQLSRGQGFQTOQTEELKRLHEEIKAKNALA 1332  
QY 466 -----ROODEDFRQLEEKGRKTAENKVMTELMTINKWRLLY-----EELYKT 511  
DB 1333 HALQSARHDCCELLREQYEEQEAKEGELQALRSKANSEVAQWRKYETDAIORTLELEAK 1392  
QY 512 KPFOQQLD-----AFEAQKQALINE-----HGAQTEQOLNKTRDSYA 547  
DB 1393 KKLQRLQDAEHEVAVNAKCALEKTKQRLQNEVEDLMVDVERNSNAACAALDKKQKQNF 1452  
QY 548 QLLGHQNLK-----VKR--KQNELRLQELGDKA 600  
DB 1453 KILAEMKQKYEETQTELEASOKESRSLSTELFKMKNAYEESLDHLETLKRNKNLQOEITA 1512  
QY 578 KIRSQ-----VKR--KQNELRLQELGDKA 600  
DB 1513 DLTEQIABEGKAVHELEKVKKHVQEKSELOASLEEA 1549

RESULT 11  
YL17\_CABEL STANDARD; PRT; 1130 AA.  
ID YL17\_CABEL  
AC Q11102;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 131.5 kDa protein C02F12.7 in chromosome X.  
GN C02F12.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Miller N.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: WEAK, TO MYOSINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U41545; AAK39135.1; -.  
DR WormPep; C02F12.7; CE03901.  
KW Hypothetical protein; Coiled coil.  
FT DOMAIN 121 779 COILED COIL (POTENTIAL).  
FT DOMAIN 805 1061 COILED COIL (POTENTIAL).  
SQ SEQUENCE 1130 AA; 131485 MW; B0FD2EFE3D99FB09 CRC64;

Query Match 13.1%; Score 408; DB 1; Length 1130;  
Best Local Similarity 24.3%; Pred. No. 1.6e-07;  
Matches 179; Conservative 132; Mismatches 270; Indels 156; Gaps 32;

QY 7 ELMKLRNKRRTKMRSMVMKQEG-----MELKLAQ--TQDLTSGKIVQLEGKLVSTKEK 59  
DB 290 ELEALRDHEEALKEALLARQDEFDRLQELQLSKRSREDLVSKNDVTALEKLLHNKEK 349  
QY 60 E-----KIDE-KCETKLEYIQEISCASDOVECKVDVIAQLEDLEKOR---EILSL 109  
DB 350 EVQTLTKELDVQKTEINDKIRRLTEVTSEFAEYRK---KFOQOEELRLKALLTVVEAA 406  
QY 110 KQSPLENTT-ESKQIEDLVKQCLETERDNLVSKDRERATLSAEMOIL-----TER 161  
DB 407 KEKLESVTSIDLOVEVKALKNKVEFLKEKERENLQSOSESQTQLOSSQVDALEAVLHSTYKE 466  
QY 162 LALEREQYEKLOQKFLQSSLIQ--QEKELSAR---LQOOLCSFOEMTSEKNVFEELK 216  
DB 162 LALEREQYEKLOQKFLQSSLIQ--QEKELSAR---LQOOLCSFOEMTSEKNVFEELK 216



```
QY 114 BENIT-FSKQIEDLVKCOLLETERDNLVSKDRAETLSAEMOILTERLALEREQEK 172
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 945 EDECSLAKDIDDLLETLAKVEKEHATENK-----VKNLTTEMASQDSIAKLTKKKAL 1000
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 QOKELQOSLSLQOKE-----LSARLOQLCSFQEDMTSEKNV-----FKBEL 215
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1001 QEAHQOITDLDQAEBDKVNTLTAKTKLEQVDDLEGSLEQEKKLMDLRAKRRKLEGD 1060
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 216 KLALAEADVQKEQESERLVKQLE-----EERKSTAEQLTRLDNLLREKEVELE 265
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1061 KLAQESIMDLENEKQSDSEKIKKKDFEISQLLSKIEDQSLGAOLKIKELQARIELE 1120
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 266 KHIAHAQAAILTAQBYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESOEKYN 325
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1121 BEIEAERAARAKVEQRADLSRELEISERLE---EAGGATA-----AQIENKKREA 1170
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 326 DTAQSLRDVTAQLESQEKYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESYKS 385
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1171 EFOKMRD-----LEESTLQHEATAAALR--REQADSAE-LGEQIDNLRVQKQLEKEKS 1223
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 386 STLKEIEDL-----KLENL--TLQEKVA-MAEKSVEDVQO-----QILTAE 424
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1224 EYKMEIDLTSMERAVAKANLEKMCRTLEDQSEIKTKSDENVRLQNDMNAORARLOT 1283
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 TNOEYARMVQDLQNRSTLKEEIKET---SSFLEKITDLKNLQOQ-----468
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1284 ENGFSRQLEE-----KEALVSQLTRGQAYTQOIEELKRHIEEVKAKNALAHAVQS 1336
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 469 ---DED-FRKQLEKGGKTAKENVMTELTWEINKWRLLY-----EELYEYTKPQQ 516
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1337 ARHDCDLLREQYEEQEAELQGMKANSVAQWRTKYETDAIQTREELAEAKKLAQ 1396
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 517 OLD-----APEAKQALLNE-----HGATQEQLN-----540
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1397 RLQDAEESIEAVNSKASLETKQKRLQGEVDELMDIVRANSANLQKQNFQKVLAE 1456
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 541 ---KIRSYAQLLGHQ-----NLKQIKHVVKLDKENSQKSEVSKLSRQ 582
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1457 WKQYEEQAELEGAQKARSLSLTFPKMKNSEYEAHLDTLKREKNLQOEISDLTEQ 1516
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 593 L-----VKKQNEURLQGLDQKALGIRHFDPSK 610
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1517 LGETGKSIHEKAKKTVESEKSE--IQTALEEAEGTLEHEESK 1558
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 13
MYH3_CHICK
ID MYH3_CHICK STANDARD; PRT; 1940 AA.
AC P02565;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87194881; PubMed=3571266;
RA Molina M.I., Kropp K.E., Gulick J., Robbins J.;
RT "The sequence of an embryonic myosin heavy chain gene and isolation
of its corresponding cDNA."
RL J. Biol. Chem. 262:6478-6488(1987).
RN [2]
RP SEQUENCE OF 1502-1940 FROM N.A.
RX MEDLINE=83161144; PubMed=6833296;
RA Kavinsky C.J., Umeda P.K., Sinha A.M., Elzinga M., Tong S.W., Zak R.,
RA Jakovic S., Rabinowitz M.;
RT "Cloned mRNA sequences for two types of embryonic myosin heavy chains
```

```
from chick skeletal muscle. I. DNA and derived amino acid sequence of
light meromyosin."
J. Biol. Chem. 258:5196-5205(1983).
-!- FUNCTION: MUSCLE CONTRACTION.
-!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
-!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
-!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
-!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
-!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
-!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; V00430; CAA23712.1; -
EMBL; J02714; AAA48972.1; -
PIR; A02990; A02990.
PIR; A29320; A29320.
HSSP; P13538; 2MVS.
InterPro; IPR000048; IQ_region.
InterPro; IPR004009; Myosin_N.
InterPro; IPR002928; Myosin_tail.
InterPro; IPR001609; myosin_head.
Pfam; PF00063; myosin_head; 1.
Pfam; PF00612; IQ; 2.
Pfam; PF01576; Myosin_tail; 1.
Pfam; PF02736; Myosin_N; 1.
PRINTS; PR00193; MYOSINHEAVY.
ProDom; PD000355; myosin_head; 1.
SMART; SM00015; IQ; 1.
SMART; SM00242; MYSC; 1.
PROSITE; PS50096; IQ; 1.
Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
Calmodulin-binding; ATP-binding; Methylation; Alkylation;
Multigene family.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.
FT DOMAIN 844 1940 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP.
FT DOMAIN 660 682 ACTIN-BINDING.
FT DOMAIN 762 776 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 700 700 ALKYLATION (SH-1).
FT MOD_RES 710 710 ALKYLATION (SH-2).
FT VARIANT 379 379 G -> D.
FT CONFLICT 1547 1547 T -> A (IN REF. 2).
FT CONFLICT 1913 1915 ERA -> GRT (IN REF. 2).
SQ SEQUENCE 1940 AA; 222816 MW; C34833D75B04DF2 CRC64;
Query Match 13.0%; Score 404; DB 1; Length 1940;
Best Local Similarity 23.2%; Pred. No. 3.7e-07;
Matches 177; Conservative 138; Mismatches 233; Indels 214; Gaps 29;
QY 9 MKLRNKRKTFMFSMMVKQF--GMELQATOKDLTFESKGIQVLEKGLVSIIEKIDKDC 66
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 835 MKLFKFKPLKLSAASEKEMANNKKEFEKTEELAKSEAKKKEKMKVSLQERNDLQ 894
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 67 ETEKLEVIQETSCASDQVEKCKV-----DIAQLEDLKEKREILSLKQSEENIT- 118
: : : : : : : : : : : : : : : : : : : : : : : : : :
```

Db 895 QVQAEADGLADABERCDQLIKTKIQLEAKIKELTERAEDEEEMNAELTAKKRKLEDCSE 954  
QY 119 FSKOIEDTLVKCOLLTERDNLVSKDRERAETLSAEMOILTERLALEROEYKLOQKEIQ 178  
Db 955 LKKDIDDLLETLAKVEKEKHATENK-----VKNLTEMALDETIAKTKKALQEAHQ 1010  
QY 179 S-----QSLQOQKEKEL-----SARLQOQKCSFQ- 202  
Db 1011 TLDDLQAEEDKVNLTAKTKLEQQVDDLEGSLEQEKLRMDLERAKRKLEGLDKTKTQES 1070  
QY 203 --PMTSEKNVFKELKALAEALDAVQOKEQSERLVKQLEERKSTAEQTLRLDNLRLRK 260  
Db 1071 TMDLENDKQOLDEKLKAFETISQKIEDEQALGMOLQKKIK---ELQARIEEL--BE 1125  
QY 261 EVELEKHIAHAQAILIAOEKYN-DPAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLES 319  
Db 1126 EIAEATSRKA-----EKHRADLSRELEISERLE---EAGGATAQI-DMNKKREA 1174  
QY 320 EQEKYNDTAQSLRDVTAQLESE-----QEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRD 375  
Db 1175 EFQKMR---RDLEEATLQHEATAAALRKHKH---ADSTADVGEQIDNLQR----- 1217  
QY 376 VSAQLESYSKSTLKETEDLKLNLTLQEKVMAEK---SVEDVQOQILTAESTNQ----- 427  
Db 1218 VKQLEKSEKSELKMEIDDLASNMESYSKAKANLEKRCRSLQSLSEIKTKEEQORTIND 1277  
QY 428 -----EYARMYQDLQNRSTLKEEIKET---SSFLEKITDLKNLQROO--- 468  
Db 1278 ISAKARLOTESGEYSRVEE-----KDALISLSRQKQAFQOIEBLKRLHEEIK 1330  
QY 469 -----DED-FRQLEEKGRTAENKENVNTELTMEINKWRLLY-----BE 506  
Db 1331 KKPCHALQASRHDCLLREQVEEQAEGELQALSKANSEVAQWRTKYETDAIQTREE 1390  
QY 507 LYEKTFPFQOOLD-----AFEAQKQALLNE-----HGATQOLNWKI 542  
Db 1391 LEEAKKLAQRLQDAEEHVEAVNSKASLEKTQORLQNEVEDLMDIVERSNAACAALDK 1450  
QY 543 RDSYAQLLGHQNLK-----KQIKHVVKLDKDNSOL 572  
Db 1451 OKNFDKILSEWKOKYBETOAELEASQESRSLSTELFKMKNAYERSLDHLETKRENKUL 1510  
QY 573 KSEVSKLRSQ-----VKR--KONELRLOGLDKA 600  
Db 1511 QOEISDLTEQIAEGGKRAIHELEKVKKQIEQKSELQTALEEA 1552  
RESULT 14  
MYSS\_RABIT STANDARD; PRT; 1084 AA.  
AC P02562;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Myosin heavy chain, skeletal muscle (fragments).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE OF 1-258.  
RA Capony J.-P., Elzinga M.;  
RT "The amino acid sequence of a 34,000 dalton fragment from S-2 of  
RT myosin."  
RL Biophys. J. 33:148A-148A(1981).  
RN [2]  
RP SEQUENCE OF 259-428.  
RX MEDLINE=85131142; PubMed=3972832;  
RA Lu R.C., Wong A.;  
RT "The amino acid sequence and stability predictions of the hinge  
RT region in myosin subfragment 2";  
RL J. Biol. Chem. 260:3456-3461(1985).  
RN [3]

RP SEQUENCE OF 409-1084 FROM N.A.  
RX MEDLINE=87304245; PubMed=3305014;  
RA Maeda K., Sczakiel G., Wittinghofer A.;  
RT "Characterization of cDNA coding for the complete light meromyosin  
RT portion of a rabbit fast skeletal muscle myosin heavy chain.";  
RL Eur. J. Biochem. 167:97-102(1987).  
CC -!- FUNCTION: MUSCLE CONTRACTION.  
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE  
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
CC SUBFRAGMENT (S2).  
CC -!- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,  
CC WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE  
CC CONSERVED.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; X05958; CAA29391.1; -.  
DR PIR; A02985; A02985.  
DR PIR; A05280; A05280.  
DR PIR; S00084; S00084.  
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW ATP-binding; Multigene family.  
FT NON\_TER 1 1 ALPHA-HELICAL TAILPIECE (S2).  
FT DOMAIN <1 >258  
FT NON\_CONS 258 259  
FT DOMAIN <259 1084  
FT DOMAIN 455 1084  
FT VARIANT 405 405 L -> V.  
FT VARIANT 408 408 V -> L.  
FT VARIANT 421 421 E -> D.  
FT VARIANT 423 423 S -> G.  
FT VARIANT 426 426 K -> R.  
SQ SEQUENCE 1084 AA; 125488 MW; 229CFD69A5E1F7F0 CRC64;  
Query Match 12.9%; Score 400.5; DB 1; Length 1084;  
Best Local Similarity 22.5%; Pred. NO. 2.8e-07;  
Matches 166; Conservative 141; Mismatches 254; Indels 177; Gaps 25;  
QY 16 ETKRSMVMVQEGMELKLOATQKDLTESGKIVQLEGKLVSTEKID----EKCTEKL 71  
Db 3 ETEKEMANKKEE-----FEKTKESLAKAAKRAKKEEKVLMQEKNDLQVQAEADSL 57  
QY 72 --LEVIOISCAQDQVEKCKVDIAQLEEDLKEKDEILSLKQSLLENIT-FSKQIEDLV 128  
Db 58 ADAERQDLIKTKQLEAKIKEVTERAEDEEINELTAKRKLEDECESEKDKDIDDL 117  
QY 129 KCOLTERDNLVSKDRERAETLSAEMOILTERLALEROEYKLOQKESLQSQLEQKE 188  
Db 118 TLAKVEKEKHATENK---VKNLTEMAGLDETIAKTKKALQEAHQQTLDLQAEED 173  
QY 189 -----LSARLQOQKCSFQEMTSEKNV-----FKELKALAEALDAVQOKEEQ 231  
Db 174 KVNLTAKTKLEQQVDDLEGSLEQEKIRMDLERAKRKLEGLDKLAQETSHNDIENQ 233  
QY 232 SERLVKOLE-----EERKSTAEQTLRLDNLRLREKEVELEKHIAHAQAILIAOEKY 282  
Db 234 LDEKLLKLEFNTNLSQKIEDEQALMTNLRQIEEL--EEIEAER--ASRAK-----EKQ 284  
QY 283 NDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESEQ 342

```
Db 285 RS-----DLSELEIEISERLEEAGGA---TSAQIEMNKKREAEPKRRD-----LEEAT 331
Qy 343 EKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVSAQLESYKSSPTLKIEDIKLENLTQ 402
Db 332 LQHEATAAALRKKHA--DSVAE-LGEQIDNQRVKQKLEKSEKSELMKWEIDDLAGNMETVS 388
Qy 403 EKVAMAESVEDVOOQILTAESTNQEYARMVQDQLQNRSTLKEEIEKIT-----451
Db 389 KAKGNLXKMCRTLEDQLSEVTKKEEHQRLNELSAQKARLTHESGESRQLDEKDMVS 448
Qy 452 -----SFLKEITDLKNQLRQ-----DED-FKQLKEEGKRTAEKEN 488
Db 449 QLSRGGAFTQOIEGLKRLQLEETKAKSALAHALQSSRRDCDLLREQYEEQEAKAELQR 508
Qy 489 VMTLETMTINKWL-----LYEELYEKTQPFQOQLD-----AFEAEEKQ 526
Db 509 AMSKANSEVSQWRKYCTDAIQTTEEEAAKKLAQRLQDAEEHVAVNSKASLEKTKQ 568
Qy 527 ALLNE-----HGATOBQLNKIRDSYAQLLGHQNLK-----556
Db 569 RLQNEAEDLMDIVERSNATCARMDKQKQNFQKVLAEWKHKYEETQAELEASQKESRSLST 628
Qy 557 -----QKIHVVVKLKDENSOLKSEVSKLSQL-----VVRK--QNELR 592
Db 629 EVFKVKNAYEESLDHLETKRENKLNQOEISDLTQOIAESAKHIHELEKVKKQIDQEKSE 688
Qy 593 LQGLDKALGIRHFDPSK 610
Db 689 LQALEEAEAGSLEHEEGK 706

RESULT 15
MYH3_HUMAN
ID MYH3_HUMAN STANDARD; PRT; 1940 AA.
AC P11055; Q15492;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic (Muscle embryonic
GN myosin heavy chain) (SMHCE).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263803; PubMed=2726495;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L.,
RA Rubinstein N.A., Kelly A.M., Sarkar S.;
RT "Nucleotide sequence of full length human embryonic myosin heavy
RT chain cDNA."
RL Nucleic Acids Res. 17:3591-3592(1989).
[2]
RP SEQUENCE OF 774-1940 FROM N.A.
RX MEDLINE=90033298; PubMed=2806546;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L.,
RA Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.;
RT "Human embryonic myosin heavy chain cDNA. Interspecies sequence
RT conservation of the myosin rod, chromosomal locus and isoform
RT specific transcription of the gene."
RL FEBS Lett. 256:21-28(1989).
[3]
RP SEQUENCE OF 856-1940 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
RT myosin heavy chains."
RL Eur. J. Biochem. 189:55-65(1990).
[4]
```

```
RP SEQUENCE OF 856-1940 FROM N.A.
RX MEDLINE=89366648; PubMed=27711643;
RA Karsch-Mizrachi I., Travis M., Blau H., Leinwand L.A.;
RT "Expression and DNA sequence analysis of a human embryonic skeletal
RT muscle myosin heavy chain gene."
RL Nucleic Acids Res. 17:6167-6179(1989).
CC -|- FUNCTION: MUSCLE CONTRACTION.
CC -|- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -|- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -|- DEVELOPMENTAL STAGE: ABUNDANTLY PRESENT IN FETAL SKELETAL MUSCLE
CC AND NOT PRESENT OR BARELY DETECTABLE IN HEART AND ADULT SKELETAL
CC MUSCLE.
CC -|- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -|- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -|- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -|- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -|- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13988; CAA32167.1; -
CC DR EMBL; X13100; CAA31492.1; -
CC DR EMBL; X51593; CAA35942.1; -
CC DR EMBL; X15696; CAA33731.1; -
CC DR PIR; S04090; S04090.
CC DR HSP; P13538; 2MYS.
CC DR Genew; HGNC:7573; MYH3.
CC DR MIN; I60720; -
CC DR InterPro; IPR000048; IQ_region.
CC DR InterPro; IPR004009; Myosin_N.
CC DR InterPro; IPR002928; Myosin_tail.
CC DR InterPro; IPR001609; myosin_head.
CC DR Pfam; PF00063; myosin_head; 1.
CC DR Pfam; PF00612; IQ; 2.
CC DR Pfam; PF01576; Myosin_tail; 1.
CC DR Pfam; PF02736; Myosin_N; 1.
CC DR PRINTS; PR00193; MYOSINHEAVY.
CC DR ProDom; PD000355; myosin_head; 1.
CC DR SMART; SM00015; IQ; 1.
CC DR SMART; SM00242; MYSC; 1.
CC DR PROSITE; PS50096; IQ; 1.
CC DR PROSITE; PS50096; IQ; 1.
CC KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IQ.
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 656 678 ACTIN-BINDING.
FT DOMAIN 758 772 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (POTENTIAL).
FT MOD_RES 696 696 ALKYLATION (SH-1).
FT MOD_RES 706 706 ALKYLATION (SH-2).
FT CONFLICT 1331 1331 A -> G (IN REF. 3).
FT CONFLICT 1391 1392 KK -> QE (IN REF. 1 AND 2).
FT CONFLICT 1609 1609 SR -> RA (IN REF. 3).
FT CONFLICT 1663 1664 RG -> QT (IN REF. 2).
FT SEQUENCE 1940 AA; 224035 MW; 43CA58G6A4BA1253 CRC64;
```



```

Query Match      12.9%; Score 400; DB 1; Length 1940;
Best Local Similarity 21.9%; Pred. No. 5e-07;
Matches 168; Conservative 137; Mismatches 261; Indels 200; Gaps 22;

QY 4 LSLLEMLKLNKRTKMRSMVMVQEGMELKLOATQKDLTSGKGIYVQEGKLYSIEKEKID 63
Db 833 LFFKIPLKLSAETKEMATKKE-----FOKTKELAKSEAKKEKLVTLVQEKND 887
QY 64 -----EKCELEYEQEISCDASDVQCKV-----DIAOLEEDLKEKREILSLKOS 112
Db 888 LOLOVQAESENLLD--AEERC--DOLIKAKFOLEAKIKEVTERAEDEEINEALTAKKRK 943
QY 113 LENIT-FSKQIEDLVKCOLLETERNLVSKDRERAETLSAEMQILTERLALERQOEYK 171
Db 944 LEDECSLKDDIDDLTILAKVEKEKHATENKVNLTTELSG----LDETIAKLTREKKA 999
QY 172 LQOKELQSOSLLOQKE-----LSARLQOQLCSFQEWTSSEKNV-----FKEE 214
Db 1000 LQAHQOALDDLQAEEDKYNLSNKTSSKLEQVEDLESSEKELRVLDLERNKRLKLEGD 1059
QY 215 LKLALAE-LDAVQOKEEQSERLVK-----QLE---EERKSTAEQLTRDNLNREKEVEL 264
Db 1060 LKLAQESILDENDKQOLDERLKKDFEYQLOSKVEDQOTLGLQFQKKIKELQARIEL 1119
QY 265 EKHIAAQAAILTAQKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVTAQLESEQE-K 323
Db 1120 EEEIEAERATRAKTEKORSYARELEBELSERLEE-----AGGVTSTQIEINLKKREAE 1171
QY 324 YNDTAQSLRDVTAQLESE-----QEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVSAQ 379
Db 1172 FLKRLRDLEEATLQHEAMVATLURKKHADSVAEIGEIDNLQRVQK-----1217
QY 380 LESYKSTLKEITEDLLENLTLOEKVAMAEKSVEDVQOQILTAESTNQEYARMVODLQNR 439
Db 1218 LEKESEFKLEIDDLSSSWESYKSKANLEKICRTLEDQLESEARGNNEIQSLSLTTO 1277
QY 440 STLKEEIEKEIT-----SSFLEKITDLKNLQRODEDFRQOLEEKGK-----482
Db 1278 KSRLOTEAGELSRQLEEKESIVQSLRSKQFTQOTTEELKQLEEBENKAKNALAHALQSS 1337
QY 483 -----TAEKENVMTLMEINKHRLLY-----EELYEKTQPFQOQ 517
Db 1338 RHDCDLLREQYEEEOEGKAELQALSKANSEVAQWRTKYETDAIQORTEEEBAKKLAOR 1397
QY 518 L-----518
Db 1398 LQDSEEQVEAVNAKASLEKTKORLOGEVEDLMVDVERANSAAALDKQORNFQVLAEW 1457
QY 519 -----DAFEAEKQALLNEHGATQQLNKIRDSYAQLLG-----HQNLKQKIKHVYK 566
Db 1458 KTKCESQAELEASLKESRSLSTELFKLNAYEALDQLETVKRENKNEQBIADLTEQI 1517
QY 567 DENSQKSEVSKLRSQLVK-----QNELRQOGELDKALGIR 604
Db 1518 AENGKTIHELEKSRQIOLEKADIQALAEAEAALEHEEAKILRIQ 1563

```

Search completed: December 20, 2002, 15:12:10  
Job time : 24.3068 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:11:28 ; Search time 7.91077 Seconds  
(without alignments)  
1337.138 Million cell updates/sec

Title: US-09-685-010-48

Perfect score: 3104

Sequence: 1 MRALSLEMLKRNKRETKMR.....FCHASKENFTPLKEGNPNCC 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	349	11.2	1958	12	US-10-028-946-4
2	349	11.2	2054	12	US-10-028-946-2
3	347	11.2	868	9	US-09-884-001-19
4	345	11.1	2053	9	US-10-017-216-2
5	342	11.0	2055	9	US-10-017-216-4
6	338.5	10.9	879	9	US-10-108-605-217
7	331	10.7	1597	9	US-10-017-216-6
8	324	10.4	1641	9	US-10-017-216-5
9	320.5	10.3	2125	10	US-09-919-172-29
10	302.5	9.9	1045	10	US-09-815-242-10617
11	302.5	9.7	2568	10	US-09-866-108-3
12	300	9.7	909	10	US-09-925-299-988
13	299.5	9.6	660	10	US-09-864-761-47959
14	289.5	9.6	677	10	US-09-745-763-168
15	299	9.6	2310	9	US-09-991-496-120
16	299	9.6	2310	10	US-09-874-923-120
17	298	9.6	2139	10	US-09-727-384-6
18	288.5	9.3	689	9	US-10-108-605-305
19	286.5	9.2	1786	9	US-09-742-096-3

20	284.5	9.2	1192	10	US-09-815-242-10903	Sequence 10903, A
21	278	9.0	751	10	US-09-864-761-38419	Sequence 38419, A
22	274.5	8.8	900	12	US-10-071-751-21	Sequence 21, Appl
23	273.5	8.8	1711	10	US-09-771-161A-219	Sequence 219, App
24	273.5	8.8	1711	10	US-09-771-161A-220	Sequence 220, App
25	267.5	8.6	2285	10	US-09-932-183A-2	Sequence 2, Appl1
26	266	8.6	1179	10	US-09-815-242-13262	Sequence 13262, A
27	265.5	8.6	945	10	US-09-745-763-191	Sequence 191, App
28	265	8.5	830	9	US-10-033-245-7	Sequence 7, Appli
29	265	8.5	830	9	US-10-033-223-7	Sequence 7, Appli
30	265	8.5	830	9	US-10-033-167-7	Sequence 7, Appli
31	265	8.5	830	12	US-10-033-246-7	Sequence 7, Appli
32	265	8.5	830	12	US-10-033-301-7	Sequence 7, Appli
33	265	8.5	830	12	US-10-033-326-7	Sequence 7, Appli
34	265	8.5	1286	9	US-10-017-216-7	Sequence 7, Appli
35	263.5	8.5	600	10	US-09-975-901-2	Sequence 2, Appli
36	258	8.3	1184	10	US-09-815-242-5229	Sequence 5229, Ap
37	258	8.3	1188	10	US-09-815-242-12125	Sequence 12125, A
38	258	8.3	1242	10	US-09-925-299-911	Sequence 911, App
39	255.5	8.2	981	10	US-09-815-242-12211	Sequence 12211, A
40	255	8.2	1179	10	US-09-815-242-13608	Sequence 13608, A
41	244.5	7.9	1048	10	US-09-741-669-409	Sequence 409, App
42	244.5	7.9	1048	10	US-09-815-242-10062	Sequence 10062, A
43	244	7.9	435	10	US-09-866-582-33	Sequence 33, Appl
44	241.5	7.8	2835	10	US-09-885-535-4	Sequence 4, Appli
45	241	7.8	996	10	US-09-815-242-5251	Sequence 5251, Ap

ALIGNMENTS

RESULT 1  
US-10-028-946-4  
; Sequence 4, Application US/10028946  
; Patent No. US20020123622A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Miranda, Carl Johan  
; APPLICANT: Friddle, Carl Johan  
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encodi  
; FILE REFERENCE: LEX-0289-USA  
; CURRENT APPLICATION NUMBER: US/10/028,946  
; PRIOR FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: US 60/258,335  
; PRIOR FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 1958  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-028-946-4

Query Match 11.2%; Score 349; DB 12; Length 1958;  
Best Local Similarity 22.3%; Pred. No. 6.3e-09;  
Matches 154; Conservative 151; Mismatches 239; Indels 148; Gaps 29;  
Qy 25 KQEGHKLQATQDLTESKGIKIVOLEKLVSI-----EKEKIDKCEK-L 71  
Db 443 KTSSEKRLIKSLKELQSDQCKHKEQEMTLRHRVSEVAVLSQKVELKASETQKSL 502  
Qy 72 LE-----YIQETSCASDOVEKCKVDIA-----OLEEDLKEKREILSLKQSLLENITF 119  
Db 503 LEQDLATVITECCSLKRLSEQARMEVSQEDDKALQLLHDIREQSKRLQEIKEQ-----EY 557  
Qy 120 SKQIEDLVKCOLLETEDNVLNVDKDRERAEITLSAEMQILTERLALER----- 166  
Db 558 QAQVEEMRLMNQLE---EDLVSA-RRRSDLYESEL--ESRLAAEFKRRKATECOHKL 611  
Qy 167 -----QYEEKLOKEQLQSLLQKELSLRLOOQLCSFOFEMTSEKNVFEKEL 216  
Db 612 KAKDQKGPVEGYAKLEKINAFQQLKIQ---ELQEKLEKAVKASTEATELLQNI-ROAKE 667



```
Db 121 SVI--EVTKGQLEVIQTQVTAKEVIOGEVRCCLKLELDTESQAE--QERDAAAQLAAQ 176
Qy 186 EKELARLQOQLCSQFEEMT-----SEKNVFEELKALAEALD-----AVQOKEQS 232
Db 177 EOEKGTALQOQAHEKEVQNLREKWEKERSWHOOELAKALESERKMELEMRUKEOOT 236
Qy 233 ERLVKQLEERKSTAEQTLRLDNLRLREKEVELEKHAHAQAAILIAQAEKYNDTAQSLRDV 292
Db 237 EMEAIAQAE-----EERTQAESALCQMLETERKERSVLSLETLTQKELADASQQLERL 291
Qy 293 TAQLESVOEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESE-----QEKYND 347
Db 292 RQDMKVQKLKEQETTGILO---TQLEAQRELKEAARQHRDDLAALQBESSLLQDKM-D 347
Qy 348 TAQSLRDVTAOL-----ESVOEKYNDTAQSLRDVSAQLESYKSSSTLKEIEDLKL 396
Db 348 LQKQVEDLKSQLVADDSQRLVEQVEKLRRET-QEYNRIOKELEERKASLTLSMEKEQ 406
Qy 397 ENLTLOEKVMAEKSVEDVQOQILTAESTNQEYARMVDQLNRLSTLKEEIEKEITSSFLE 456
Db 407 RLLVLQEADSIQOELSALRQDMQEAQGEQKELSAQE-----LLRQEVKEKEADFLA 459
Qy 457 KITDLKNLQRODEDFRQLEEKGRKTAENKVMTELWEINKW--RLLYEELYEKTQPF 514
Db 460 QEAQALLEL-----EASHITEQQLRASLWAQEAQAAQOLQLRLRST 499
Qy 515 QOOLDFAFEAKQ-----ALLNEHGATQOLNKRIRDSYAQLLG----- 551
Db 500 ESQLEALAEQPGNQAAQALASLYSALQALGVCSERPELSGGGDSAPSVMGLBPD 559
Qy 552 -----HQNLD--KQIKHWV-----KLKD-----E 568
Db 560 QNGARSLFKRGPLLTALSAAVASALHKLHQDLWKQTQTRDVLRDQVQKLEBLTDTEAE 619
Qy 569 NSQLKSEVSKRSQLVKR-----KONELR-----LQSELDKA-LGIRHF 606
Db 620 KSVQHTQLQDLQRLQSONQEEKSKWEKQNSLESELMELHETMASIQSLRRLRAELQMEA 679
Qy 607 DPKAFCHASKENFT 621
Db 680 QGERELLQAANKLT 694
```

```
RESULT 4
US-10-017-216-2
; Sequence 2, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prot
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR FILING DATE: 2001-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-216-2
```

```
Query Match 11.1%; Score 345; DB 9; Length 2053;
Best Local Similarity 22.4%; Pred. No. 1e-08;
Matches 152; Conservative 154; Mismatches 236; Indels 136; Gaps 29;
```

```
Qy 25 KOEGMELKQATQKDLTSEKQKIVOLEGKLVSI-----EKEKIDKCEK-L 71
Db 443 KTSMEKKLLIKSKELOQSDQKCHKMEQEMTLRHRVSEAVLSQKEVELKASETQRS 502
```

```
Qy 72 LE-----YIQEISCASDQVEKCKVDIA-----QLEEDLKEKDRILSKOSLEENITF 119
Db 503 LEQDLATYITECCSLKRSLEQARMEVSDDDKALQLLHDIRQSRKLOIKEQ-----EY 557
Qy 120 SKQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQIILTERLALB--ROEYEEKLOOKEL 177
Db 558 QAQVEEMRLMNQLE---EDLVS-A-RRSDLYESEL--ESRLAAEFERKATECOHKLL 611
Qy 178 OSQSLLQOB-----KELSARLQOQLCSFOREMTSEKNVFEELKALAEALDAVQOKEE 230
Db 612 KAKQDGKPEVGYAKLEKINAEOQLKIQLQEKLE-----KAKEARELEKLNQRED 665
Qy 231 QSERLVKQI---EERKSTAEQTLRLDNLRLREKEVELEKHAHAQAAILIAQAEKYNDTAQ 287
Db 666 SSEGIRKLLVEAEERRHSLKLVKRLTWER-RENKLDKDDIQTKSQOI-----Q 713
Qy 288 SURDVTQLESVOEKYNDTAQSLRDVTAQLESQEKY-----NDTAQSLRD---VTA 336
Db 714 QMADKILELE---EKHREAQVSAQHLEVHLKQEQHYEEKIKYVDNQIKKDLADKETLEN 770
Qy 337 QLESQEKYNDTAQSLRDVTAQLESVOEKYNDTAQSLRDVSAQLESYKSSSTLKEIEDLKL 396
Db 771 MQRHEEAHAKGKILSEOKAMINAMDSKIRSEQLRIVELSEANKLAANSSSLFTQRMKA 830
Qy 397 ENLTLOEKVMAEKSVEDVQOQILTAESTNQEYARMVDQLNRLSTLKEEIEKEIT---SS 453
Db 831 Q-----EEMISELRQCKFYLETQAGKL-----EAQNRKL--EEQLEKISHQDHS 872
Qy 454 FLEKTYDLKNQUR-----QODEDFRQLEEKGRKTAENKVMTEL-----TWEINKWR 501
Db 873 DKNRLLETRLREVSLEHEEQKLELQUTELQLSQERESQTLQAAARALE--SOLR 931
Qy 502 LLYEELYEKTQFOOLDFAFEAKQALLNEHGA-----TOEQLANKIRDSYAQLLG 551
Db 932 QAKTELEETTAEEETQALTAHRDEIQRRFDALRNSCTVITDLEQLNQLTEDNAE-LN 990
Qy 552 HQN--LKQIKHWVKLVKLDENSQKSEVSKRSQSLRQKQNELRQGLQELDKALGIRHFDPS 609
Db 991 NQNFVLSKQLDASGANGDEIVLRSEVDHLRREITEREM-QLTSQKQTMEAL----- 1041
Qy 610 KAFCHASKENFTPLKEGN 627
Db 1042 KTTCTMLEEQVMDLEALN 1059
```

## RESULT 5

```
US-10-017-216-4
; Sequence 4, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type P
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR FILING DATE: 2001-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2055
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-4
```

```
Query Match 11.0%; Score 342; DB 9; Length 2055;
Best Local Similarity 22.0%; Pred. No. 1.4e-08;
Matches 152; Conservative 152; Mismatches 240; Indels 148; Gaps 29;
```

```
Qy 25 KOEGMELKQATQKDLTSEKQKIVOLEGKLVSI-----EKEKIDKCEK-L 71
Db 442 KYSSMEKKLLIKSKELOQSDQKCHKMEQEMTLRHRVSEAVLSQKEVELKASETQRS 501
```



```

Qy 7 ELMLRNKRETKMRSMVKQKGMELKQATQKDLTSGKIVQLEGLVSIKEKIDK 66
Db 12 EMTRL-HRRVSEAVLSQKE-VELKASQTSRLE-----45
Qy 67 ETEKLEIYQISCASDOVECKVDIA-----OLEEDLEKEDREILSLKQSL 119
Db 46 --ODLATYITECSSLRSLQARMEVSQEDDKALQLLHDIRQSRKLOEIKEQ 98
Qy 120 SKQEDLVTKCOLLETERDNLVSDRRAETLSAEMOILTERLALER-----166
Db 99 QAQVEMLMNNQLE---EDLYSA-RRSDLYESEL--ESRLAAEFKRVKANEQKHL 152
Qy 167 -----QYEKLOQKELQSLQKELQKELQKELQKELQKELQKELQKELQK 216
Db 153 KAKDQKPEVGEYSKLEKINAEQOLKIQ---ELQEKLEKAVKASTEATELLQNI-RQAKE 208
Qy 217 LALAELEDAVQKQESQERLYKQ---BEERKSTAEQTLRLDNLLEKEVELEKHAHAQ 273
Db 209 RAERELEKLNHREDSSEGIKKLVAEAEERHSHLENKVRLETMER-RENRLKDDIOTKSE 267
Qy 274 AILIAQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKY-----N 325
Db 268 QI-----QOMADKILILE---EKHREAQVSAQHLVHLKQKQHQHYEEKIKVLDN 313
Qy 326 DTAQSLRD---VTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVSAQLES 382
Db 314 QIKKDLADKESLENMORHEEAEHEKGKILSEQKAMINAMDSKIRSLRQIRIVELSEANKL 373
Qy 383 YKSTLKEIEDLKLNTLQEKVAMAEKSVEDVQOILTAESTNQBYARMVODLQNRSTL 442
Db 374 AANSLSFTQRMNKAQ-----EMISELRQKQFYLETOAGKL-----EAQNRKL- 416
Qy 443 KEEIKEIT---SFLKIDTLKNQLR-----QODEDFRQKLEKGRKTAENVMTE 492
Db 417 -EQLEKISHQDHQSKSLLETLRLREVSLSEHEEQKLELRQJTELQLSQERESQUTA 475
Qy 493 L-----TMEINKWRLLYEYEKTPFQOQDFAEAEKQALLNEHGA-----TQE 537
Db 476 LQAAARALE-SQLRQAQTELETTAEAEETQALTAHDEIQRFDALRNSCTVITDLEE 534
Qy 538 QLNKIRSYAQLLGHQ--LKQIKHVVKLDKNSQLKSEVSKLSQLVKKRQKQNELRQOG 595
Db 535 QLNQLTEDNAE-LNNQNFYLSQKLEDEASGANDIEVLARSEVDHLRREITEREM-QLTSQK 592
Qy 596 ELKALGIRHFDPSKAFCHASKENFTPLKEGN 627
Db 593 QTMEL-----KTTCTMLEEQVLDLEALN 616

```

RESULT 8

```

US-10-017-216-5
; Sequence 5, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR FILING DATE: 2001-10-23
; PRIOR FILING DATE: 2001-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 5
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-5

```

Query Match 10.4%; Score 324; DB 9; Length 1641;  
 Best Local Similarity 22.0%; Pred. No. 7e-08;  
 Matches 148; Conservative 130; Mismatches 253; Indels 141; Gaps 25;

```

Qy 25 KOEGMELKQATQKDLTSGKIVQLEGLVSI-----EKEKIDKCEK-L 71
Db 70 KYSSMEKKLLLSKRELQSDQCKHKEQEMTLHRRVSEAVLSQKEVELKASQTSRSL 129
Qy 72 LE-----YQIEISCASDOVECKVDIA-----OLEEDLEKEDREILSLKQ-----111
Db 130 LEQDLATYITECSSLRSLQARMEVSQEDDKALQLLHDIRQSRKLOEIKEQEQVAAQVE 189
Qy 112 -----SLEENITPSQIODEL-----TVKCOLLETERDNLVSKDR-----145
Db 190 EMRLMNNQLEEDLVSAARRSDLYESELRESRLAAEFKRVKANEQKHLKMAKDOQKPEVG 249
Qy 146 --ERAETLSAEMOILTERLALERQYKELQKELQKELQKELQKELQKELQKELQK 203
Db 250 EYSKLEKINAQOLKIQEL---OEKLEKAVKASTEATELLQNIQAKERAERELEKHLNR 306
Qy 204 MTSEKNVFK-----EEL--KLALAEIDA-----VOQKEQESERLVKOLEEERKSTAEOL 250
Db 307 EDSSEGIKKLVAEAELEKHEKREAOVSAQHLVHLKQKQHQHYEEKIKVLDNQIKKDLADK 366
Qy 251 TELDNLLEKEVELEKHAHAQAAILIAQEK--YNDTAQSLRDVTAQLESQEKYNDTAQ 308
Db 367 ESELENMORHEE-----AHEKGKILSEQKAMINAMDSKIRSLRQIRIVELSEANKLAAN 420
Qy 309 SLRDVTAQLESQEKYNDTAQ--LRDVTQLESQEKYNDTAQSLRDVTAQLESQEK 365
Db 421 SSLFTQRMNKAQEMISELRQKQFYLETOAGKLEAQNRLKLE---EQLEKISHQDHQSKSR 477
Qy 366 YNDTAQSLRDVSAQLESYKSTLKEIEDLKLNTLQEKVAMAEKSVEDVQOILTAEST 425
Db 478 LLELETRLEVSLEHEEQKLELRQJTELQ---LSQER-----ESQLTALQAARAALES- 529
Qy 426 NQBYARMVODLQNRSTLKEEIEKEITS-----SFLKIDTLKNQLRQODED- 471
Db 530 --QLRQAQTELETTAEAEETQALTAHDEIQRFDALRNSCTVITDLEEOLNLTEDN 587
Qy 472 -----FRKLEKGRKTAENVMTELTMEINKWRLLYEYEKTPFQOQDFAE 522
Db 588 AELNNQNFYLSQKLEDEASGANDIEVLARSEVDHLR--REITEREMQLTSQKQVWE 640
Qy 523 AEKQ--ALLNEHGATQEQNLKIRDSYAQLLGHQNLKQIKHVVKLDKNSQLKSEVSKLSR 580
Db 641 ALKTTCTMLEEQVLDLEALN-----ELLEKERQWEAWRSV--LGDEKSQFECRVRELEQ 692
Qy 581 SOLVKRKQNELR 592
Db 693 RMLDTEKQSRAR 704

```

RESULT 9

```

US-09-919-172-29
; Sequence 29, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR FILING DATE: 2001-07-30
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 2125
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 3774181CD1
US-09-919-172-29

```

```
Query Match      10.3%; Score 320.5; DB 10; Length 2125;
Best Local Similarity 22.5%; Pred. No. 1.3e-07;
Matches 151; Conservative 141; Mismatches 253; Indels 125; Gaps 28;

Qy 7  ELMKLRNRYKRMVMYKQEGMELKQATQKDLTSGKIVQLEGGKLVSTEKEDK 66
Db 636  ELPYREAENELRQQRVEDISL-----QKIRASEAK--QYRELETIVREK--EAA 686
Qy 67  EFKEKLVYQETSCASDVQEKVKVDIAOLEEDLKEKDRILSLKOSLENTFFSKOIEDL 126
Db 687  ERE--LERVQLTIFAERK-----AAVENL-----LNFRLQLENTFFRTLEDH 731
Qy 127  TVKCOL-----LETERDNLVS---KDRERAETLSAEMQILTERLALEREYKLOQKEQLS 179
Db 732  LKRKDLNDLQEQKMLMEELRRKRDNEELLKLIKQMEKDLAFQKQVAEK-QLKE--- 787
Qy 180  QSLQOEKELSRLQ---QQLC-----SQEEMTEKKNVFEEL 215
Db 788  ----KOKIEARRKITEIQTCTRENALPVPITQATSCRAVTGLQOEHDQK-----ABEL 840
Qy 216  KLALAEIDAQOQKQSER-LVKOL-----EERKSTAEOLTRLDNLLREKEVEL 264
Db 841  KOQVDELTAANKKADQDRELTYELNALOLEKTSSEKARLLKDKLDNETNWLRLCKLEL 900
Qy 265  EKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESQEKYKYNDAQSLRDVTAQLES---EQ 321
Db 901  ER---KQAEKGYSQLRELGRQLNQTTGAEEAMQASDLKIKRNVQLESLNHEK 956
Qy 322  EKYNDTAQSLRDVTAQLESQEKYKYNDAQSLRDVTAQLESVQ--EKYNDTAQSLRDVSAQ 379
Db 957  GKLOREVDRIITRAHVAEKNIOHLNSQIHSFRD-EKELERLQICQKSD-----HLKEQ 1009
Qy 380  LESYKSSYTLKETEDLKLNLTLQ---EKVAMAEKSVEDVQOQILTAESTNOEYARMVDL 436
Db 1010  FEKSEHQLLONIKAKENNDKIQRNEELEKSNCEAEMLKQKVELTQNNETKLMQRI 1069
Qy 437  QNRS---TLKEEETKEITSSFLEKTTIDLKNOLROODE-----DFRKOLEEKGRKT 483
Db 1070  QAESENVILEKQTIQORCEALKIQADGFKQOLRSTNEHLHKQTKTEQDFQRKIKCLEEDL 1129
Qy 484  AKEKNVMTLMEINKWRLLYEELYEKTPFOOOLDAFEAK-----QALLNEHGATQE-- 537
Db 1130  AKSQNLVSEFKQKQOQOQIITQNTKTHVRNLAELNASKEERKRGQKVLQQAQVQELN 1189
Qy 538  -OLNKIRDSYAOILLGHQNLKQIKH--VVKLDENSQKSEVSKLSQLVKRRKQNELRLQ 594
Db 1190  NPLKVQDE---LHLKTEEQMTHRKVLPQESGKFKQSAEERPKMKELMESKVITE 1245
Qy 595  GELDKALGIR 604
Db 1246  NDIS---GIR 1252

RESULT 10
US-09-815-242-10617
; Sequence 10617, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: zyskind, Judith W.
; APPLICANT: wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10617
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-815-242-10617

Query Match      9.9%; Score 306.5; DB 10; Length 1045;
Best Local Similarity 21.0%; Pred. No. 2.7e-07;
Matches 156; Conservative 148; Mismatches 255; Indels 185; Gaps 28;

Qy 7  ELMKLRN---KRETKRMSVMYKQ-EGMELKQATQKDLTSGKIVQL----- 50
Db 98  ELFKRGGTGVRKQSAKISLIVKDAQGKELRSYSKRREVDLIQELHLNANOFAQIVLIP 157
Qy 51  EGKLVISIEKIDCEKTEKLEKLEYIOEISCASDVQEKVKVDIAOLEEDLKEKDRILSLK 110
Db 158  QGEFTFIIIAKSDK---EKVLRNL-----FGTELYOLFSENLEKRLKIANQEIQTQ 207
Qy 111  OSLE---ENITFSKOIE---DLTVKCOLLETERDNLVSK---DREERAETLSAEMQ----- 156
Db 208  QKIELKTLQHLHSEPEPTMTLFEKLQLELSEQOQAEQALLVERQIATLQAKQAKQEV 267
Qy 157  ---ILTERLALEREYKELQO--KELOSLSLQOEKELSARL-----QQOLCSFQEMTSE 207
Db 268  RYAIERONLQOQKPELLEKRAKQVEQTVIERLEQIQQLKWSQKQSLAEKVEKRS 327
Qy 208  KNVFKEELKALAEIDAVQOKEEQSERLVKOLEEERKSTAEOLTRLDNLLREKEVELEKH 267
Db 328  KOQKEQTSKQOALMETQOALTDMQAIMSELEEQPLIAEKQERLQTIQRO----- 379
Qy 268  IAAHAQAAILIAQEKYNDTAQSLRDVTAQLESQEKYKYNDAQ--SLRDVTA-----QLE 318
Db 380  -----LPQYQVEYQLAQOQIAEQANYQAIQKEYESCQKEKITLADKAVAKQFIEQ 431
Qy 319  SEQEKYNDTAQSLRDVTAQLESQEKYKYNDAQSLRDVTAQLESQEKYKYNDAQSLRDVSA 378
Db 432  GTLEKANFECSSVADHWQNFVERWQKNQKAWQKISQNOVELHQLTQRFAVEQOQKSAEPA 491
Qy 379  QLESYKSSYTLKEIEDLKLNLTL-----QEKVAMAEKSVEDVQO 417
Db 492  KLQTKKS-----QWASLQIQRLSLLLEEGEPCPVCSLSHPKQOOTHQEVSLSE--IDQAE 545
Qy 418  OILTAESTNQYARMVDL-----QNRSTLKE-----EIKEITSSFLE----- 456
Db 546  ELTEVEKTVQRTETLSALGAEKQKESQLOQOEAAAYTEEQOLAAQADLQPLTLGLTF 605
Qy 457  -----KITDLKNLRQODEDFRKOLEE---KGKRTAKENVMTELTMWEINKWR----- 501
Db 606  SOVTPAETAEVESQLAKEKQIAQKLTETSVEKORLAELEEQVAENSORFVLRQOQVETM 665
Qy 502  -----LTYEELYEYKTPFOOQOLDAFEAEKQALLNEHGATQEQOL 539
Db 666  QOSLERITIQOQMIASQLLDATVTYEMTKQOALLQOEELSAFEROKENVTTO-GETLKR 724
Qy 540  NKIRDSYAQLLGH-----QNLKQIKHVVK-----LKDENSQKSEVSKLSRQ- 582
Db 725  EMILES---TLTHLEKEQOTLQOTVAQLESQNLNAVLTQGVTEQDQLTWELKVEPTLESQ 781
```



```

:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (47)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (48)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (52)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (58)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (62)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (125)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (632)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (851)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
:
US-09-925-299-988

```

Query Match 9.7%; Score 300; DB 10; Length 909;  
Best Local Similarity 21.9%; Pred. No. 4.7e-07;  
Matches 146; Conservative 127; Mismatches 232; Indels 152; Gaps 26;

[illegible]

```

Db  829  -----LEEIVEKLKGELESSQVREHTXHLAELEK-----HMAAASAECONYAK  873
Qy  619  NFTPLKE 625
      ::
Db  874  EVAGLRQ 880

RESULT 13
US-09-864-761-47959
; Sequence 47959, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
; FILE REFERENCE: Aeoim1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47959
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AJ010770.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: SWTSSPROT HIT: Q99323, EVALUE 3.00e-17
; OTHER INFORMATION: EST_HUMAN HIT: AU132932.1, EVALUE 1.00e-105
US-09-864-761-47959

Query Match          9.6%;   Score 299.5;   DB 10;   Length 660;
Best Local Similarity 23.0%;   Pred. No. 3.6e-07;

```

Query Match 9.6%; Score 299.5; DB 10; Length 660;  
Best Local Similarity 23.0%; Pred. NO. 3.6e-07;



Matches 138; Conservative 124; Mismatches 218; Indels 121; Gaps 26;

Qy 5 SLEMLKLNKRETKRMVMVK--QEGMELKQATQKDLTESKG-KIVQLEGKLV-----56  
Db 90 TVELOQRNHHKDSQFETDIVQRMQETQRLKLEQRAELDEMYGQOIVQMKGOLIRQHMAQ 149  
Qy 57 IEKEKIDEKCEKLEKLEYIQBISCASQVQKCKVDIAQLEEDL-----KEKREIISLK 110  
Db 150 MEEMKTRHKGEMALRSYNTVNEQIKLMVAINELNKLQDTNSQKKEELGL- 208  
Qy 111 QSLSENITFSKQIIDLTVKCOLLETRDNLSKDRERAETLSAEOILTEALEROEYE 170  
Db 209 -ILEBKALQRLQEDLV---BELSFSRQI-----QARQTIARQE---SKLNEAHKSL 256  
Qy 171 KLOQKELQOSLQKQKELSLARLQOQLCSFO---EEMTSEKNVFKPEEL-KIALAEALDAVQ 226  
Db 257 TVEDLKAEIVSASESRKELEKHEAEVNTYKIKLEMLEKEKNVAILDRMAESQEAELRLR 316  
Qy 227 QK-----EQSERLVKQLEPERKSTAQQL-----TRLDNLLRKEVELE-----K 266  
Db 317 TQLLFSHEEELSKLEDELEIHRINIEKLDNGLGHIYKQQIDGLQNMESQKIETMQFEKD 376  
Qy 267 HIAAQAQALIAQKYNDAQSL-----RDVTAQLESVQEKYNDTAQSLRDVTAQLESEQ 321  
Db 377 NLIQKQILILEISKLQDLQSLVNSKSEMTLQINELQKEI-----EILQKEWEK 428  
Qy 322 EKYNDTAQSLRDVTAQLESE-QEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRD-----375  
Db 429 GTLEQEQVQLKLTLEKQMKEND---LQEKFAQLEAENSILKDEKKTLEDMLKIH 484  
Qy 376 --VSAQ-----LESYKST-----LKEIEDLKLENLFLQEKV-----AMAE 409  
Db 485 TPVQSEERLIPLDSIKSKSDSVWEKEIEIIEENEDLKQOCIQNBEIEQRNFTFSAE 544  
Qy 410 KSVEDVQOILTAESTNOEYA---RMVDLQNRSTLKEEIEKETSFSLEKI-----458  
Db 545 KNFEVNTQEL-----QEYACLLKVKDDLEDNKNQKQLEYKSKLKAUNELHLQRIIPT 598  
Qy 459 -TDILKNLQROODEFRKQLEKGRKRTAEKENVMTELTMEINKRLLYBELYEKTKPFOQQ 517  
Db 599 TVMKSSVFEDKTFVAETLEMGE-VVEKDTTELMEKLEVTKREKL--ELSQRLSDLSEQ 655  
Qy 518 L 518  
Db 656 L 656

RESULT 14

US-09-745-763-168  
; Sequence 168, Application US/09745763  
; Patent No. US20020065394A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John M.  
; LaVallie, Edward R.  
; Collins-Racie, Lisa A.  
; Evans, Cheryl  
; Merberg, David  
; Treacy, Maurice  
; Spaulding, Vikki  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; ENCODING THEM  
; NUMBER OF SEQUENCES: 219  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/745,763  
; FILING DATE: 18-Jun-2000  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sprunger, Suzanne A.  
; REGISTRATION NUMBER: 41,323  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8284  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 168:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 677 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 168:  
US-09-745-763-168

Query Match 9.6%; Score 299.5; DB 10; Length 677;  
Best Local Similarity 20.6%; Pred. No. 3.7e-07;  
Matches 145; Conservative 153; Mismatches 224; Indels 181; Gaps 29;  
Qy 24 VKQEGMELKQATQKDLTESKGKIV-QLEGKLVISKEKIDKCEKTELLEYIQBISCAS 82  
Db 1 MKNHYVPLKV---SDMKKSHDAIIDLNRKLLDVTQYTEKKLEMEKL-----LEN 50  
Qy 83 DOVEKCKVDIAQLEEDL---KEKREIISLKQSLSENITFSKQIEDLVTKC-----130  
Db 51 DLSLK---DVSRLFTVPVPEKHEKEITALKSNIVE---LKKQLSELAKKCGEDQEKIHA 104  
Qy 131 -----QLLETERONLVSKDRE-----RAETLSAE-WQILTE 160  
Db 105 LTSNTNLKMMNSOYVPVKTHEEVKMTLNDTLAKTNRELLDVKKRFEDINQEFVKIKDK 164  
Qy 161 RLALERQ-----EYKIQKQELQSQSLQKELKSARLQOQLCSF---QEMETSE 207  
Db 165 NEILKRNLENTQNIKAIEYISLAHEAKMSSLQSMRKVQDSNAELANRYKQGEIIVT- 223  
Qy 208 KNVFKELKLALAEIDAVQ-----OKEEOS 232  
Db 224 ---LHAETKAQKKELDITQECIKVKYAPIVSPFECERKFKATEKELKDQSLSEQTOKYSVS 280  
Qy 233 ERLVKQLEERKSTAEQTLRLDNLRLREKEVELEKIHAAHAQAILIAQEKYNDTAQSLRDV 292  
Db 281 EEEVKNKQENDKKEIKFTLQKDLRDKTVLIEK-----SHEMERALSRTDEL 329  
Qy 293 TAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV-----TAQLESVQEKY 345  
Db 330 NKQLKDLQSKYTE---VKNVKEKLVEENAKQTSILAVQNLQKHVPLEQVEALKKSL 385  
Qy 346 NDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESKYSTLKIEDLKLENLFLQEKV 405  
Db 386 NGTIENLKE---ELKSMQRCYEKEQQTVTKLHQLLENOKNSSVPLAEHLQIKE-AFEKVV 441  
Qy 406 AMAEKSVEDVQOQIILTAESTNOEYARMVDLQNL-RSTLKEEIEKEI-----TSSFLE 456  
Db 442 GLIKASLREKEEEE---SQNMKEEVSKLOSEVQNTQKALKKLETREVVDLSKYKATKSDLE 498  
Qy 457 -KITDLKNQL-----RQODEDFRKQLEEKGRKRTA---EKENVMTELTMET-----NKWR 501  
Db 499 TQISSILEKLANLNRYKEVEEVLHAKKEISAKDENELLHFSIEQEKDKQKCRDKSL 558  
Qy 502 LLYEELYEYKTPFOQLDAFAEAKQALLNEHGATQELQNLKIRDSYAQLL---GHONLKQK 558  
Db 559 TTITELQRRIOESAQIEAKONKITELND-----VERLKQALNGLSQTLYTSGNPTKROS 614  
Qy 559 -----IKHVVK-----LKDENSQKSEVSKLSQVKKRQKQNEL 591  
Db 615 QLIDTLQHVQKSLSEQQLADADRQHQEVIAIYRTHLLSAAQGHM 657

## RESULT 15

US-09-991-496-120  
; Sequence 120, Application US/09991496  
; Patent No. US40020169285A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Webb, John R.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Skelky, Yasir A.W.  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Coler, Rhea  
; APPLICANT: Probst, Peter  
; APPLICANT: Brannon, Mark  
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
; FILE REFERENCE: 210121.420C9  
; CURRENT APPLICATION NUMBER: US/09/991,496  
; CURRENT FILING DATE: 2001-11-20  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 120  
; LENGTH: 2310  
; TYPE: PRT  
; ORGANISM: Leishmania major and chagasi  
US-09-991-496-120

## Query Match 9.6%; Score 299; DB 9; Length 2310;

Best Local Similarity 20.6%; Pred. No. 1.3e-06;

Matches 149; Conservative 146; Mismatches 274; Indels 156; Gaps 26;

QY 13 NKRETKMRSMVMYKQEGMELKQAT-----QKDLTSEKGIQVLECKLVSIEKEID 63  
DB 557 NAEELQORLDTATQORAELEAQAARLAANAELQORLDTATQORAELEAQAARLAADRDE 616  
QY 64 EKCETEKLELYIQE-ISCASDQVECKVDIAOLEEDLEKREILSLKOSLEENI-TFSK 121  
DB 617 ARQOLAANAELQORLDTATQORAELEAQAARLAADGDEARQOLAANAELQORLDTATQ 676  
QY 122 QIEDLVKQCLLETERDNLVSKDRRAETLS-----AEMQILTERLALERQE--- 168  
DB 677 QRAELEAQAARLAADRDEARQOLAANAELQORLDTATQORAELEAQAARLAADRDEARQ 736  
QY 169 -----YEKLOQK-----ELQSQ---SLLOQEK 187  
DB 737 QLAANAELQORLDTATQORAELEAQAARLAADGDEARQOLAANAELQORLDTATQORA 796  
QY 188 ELSARL-----QQQLCSFQEMTSEKNVF---KEELKLALAE-----DAVQOKE 229  
DB 797 ELEAQAARLAADRDEARQOLAANAELQORLDTATQORAELEAQAARLAADGDEARQOLA 856  
QY 230 EOSERLVKOLEERKSTABQLTRNLNLEKEVELEKHIAHAQAAILIAQEKYNDTAOSL 289  
DB 857 ANAEELQORLDTATQORAELEAQAARLAADRDEARQOLAANAEL---QORLDTATQOR 912  
QY 290 RDVTAQLESVQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESQEKY-NDT 348  
DB 913 AELEAQAARLAADRDEARQOLAANAELQORLDTATQORAELEAQAARLAADG 965  
QY 349 AQSLRDVTAQLESVQEKYNDTAQSLRDVSAQL-----ESYKSTLKEIEDL---KLENL 399  
DB 966 DEARQOLAANAELQORLDTATQORAELEAQAARLAADRDEARQOLAANAELQORLDTA 1025  
QY 400 T-----LQEKVAMAKSVEDVQOQIILTA-----ESTNOEYARMVQDL 436  
DB 1026 TQORAELEAQAARLAANAELQORLDTATQORAELEAQAARLAADRDEARQOLAANAEL 1085  
QY 437 QN-----RSTLKEEKEIKETTSFLEKTLTKNQ---LRQODEDFRKOL 476  
DB 1086 QORLDTATQORAELEAQAARLAANAELQORLDTATQORAELEAQAARLAANAELQORL 1145

QY 477 EEKGKRTAEKENVMTELTMETINKRWRLLYEELYEYKTKPFOOQLDAF---BAEQOALLNEHG 533  
DB 1146 DTATQORAELEAQAARLAADRDEAR---QOLAANAELQORLDTATQORAELEAQAARLA 1202  
QY 534 ATQEQNLKTRDSYAQLLGHQNLKQKIKHVVKLKDE-NSOLKSEVSKLSQLYKRRKQNELR 592  
DB 1203 ANAEELQORLDTATQ---QRAELEAQAARLAADRDEARQOLAANAELQORLDTATQORAE 1260  
QY 593 LOGEL 597  
DB 1261 LEAQV 1265

Search completed: December 20, 2002, 15:14:50  
Job time : 17.9108 secs



```
QY 61 KIDKCEKLEKLEYIQEISCASDQVEKCKVDIAQLEEDLKEKDRILSLKQSEENITFS 120
DB 61 KIDKCEKLEKLEYIQEISCASDQVEKCKVDIAQLEEDLKEKDRILSLKQSEENITFS 120
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKIQKQELQSQ 180
DB 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKIQKQELQSQ 180
QY 181 SLLOQKELSLRLOOQLCSFQEEEMTSEKNVFEELKLALAEALDAVQOEQSERLVKOLE 240
DB 181 SLLOQKELSLRLOOQLCSFQEEEMTSEKNVFEELKLALAEALDAVQOEQSERLVKOLE 240
QY 241 EERKSTAEQLTRDLNLLREKEVELEKHHAAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300
DB 241 EERKSTAEQLTRDLNLLREKEVELEKHHAAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
DB 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
QY 361 SVQEKYNDTAQSLRDVSAQLESYKSTLKEIEDLKLLENLTLOEKVAMAEKSVEDVQOQIL 420
DB 361 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLLENLTLOEKVAMAEKSVEDVQOQIL 420
QY 421 TAESTNQYARMVDLQNRSTLKEEIEITSSFFLEKITDLKNLQROQDEDFRKOLEBK 480
DB 421 TAESTNQYARMVDLQNRSTLKEEIEITSSFFLEKITDLKNLQROQDEDFRKOLEBK 480
QY 481 KRTAEKENVMTELTMWINKWRLLYBELYEKTPFOOQLDAFAEKQALLNEHGATQEQLN 540
DB 481 KRTAEKENVMTELTMWINKWRLLYBELYEKTPFOOQLDAFAEKQALLNEHGATQEQLN 540
QY 541 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLRSQLVKRRKQNELRLQGELOKA 600
DB 541 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLRSQLVKRRKQNELRLQGELOKA 600
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
DB 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631

RESULT 2
US-08-477-831C-2
; Sequence 2, Application US/08477831C
; Patent No. 6429291
; GENERAL INFORMATION:
; APPLICANT: TURLEY, EVA A.
; APPLICANT: SHUWEN, ZHANG
; APPLICANT: ENTWISTLE, JOYCELYN
; TITLE OF INVENTION: HVALURONAN RECEPTOR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10020-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Rel. #1.0, ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477.831C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PIERRE, MARGARET A.
; REGISTRATION NUMBER: 30,709
; REFERENCE/DOCKET NUMBER: SIM-10
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: /desc = "RHAMM I protein"
US-08-477-831C-2

Query Match 95.1%; Score 2952.5; DB 4; Length 606;
Best Local Similarity 95.7%; Pred. NO. 3.7e-171;
Matches 604; Conservative 2; Mismatches 0; Indels 25; Gaps 1;

QY 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLQATQKDLTESKGIKIQLEGLVSIEKE 60
DB 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLQATQKDLTESKGIKIQLEGLK----- 54
QY 61 KIDKCEKLEKLEYIQEISCASDQVEKCKVDIAQLEEDLKEKDRILSLKQSEENITFS 120
DB 55 -----CASDQVEKCKVDIAQLEEDLKEKDRILSLKQSEENITFS 95
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKIQKQELQSQ 180
DB 96 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYEKIQKQELQSQ 155
QY 181 SLLOQKELSLRLOOQLCSFQEEEMTSEKNVFEELKLALAEALDAVQOEQSERLVKOLE 240
DB 156 SLLOQKELSLRLOOQLCSFQEEEMTSEKNVFEELKLALAEALDAVQOEQSERLVKOLE 215
QY 241 EERKSTAEQLTRDLNLLREKEVELEKHHAAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300
DB 216 EERKSTAEQLTRDLNLLREKEVELEKHHAAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 275
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360
DB 276 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 335
QY 361 SVQEKYNDTAQSLRDVSAQLESYKSTLKEIEDLKLLENLTLOEKVAMAEKSVEDVQOQIL 420
DB 336 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLLENLTLOEKVAMAEKSVEDVQOQIL 395
QY 421 TAESTNQYARMVDLQNRSTLKEEIEITSSFFLEKITDLKNLQROQDEDFRKOLEBK 480
DB 396 TAESTNQYARMVDLQNRSTLKEEIEITSSFFLEKITDLKNLQROQDEDFRKOLEBK 455
QY 481 KRTAEKENVMTELTMWINKWRLLYBELYEKTPFOOQLDAFAEKQALLNEHGATQEQLN 540
DB 456 KRTAEKENVMTELTMWINKWRLLYBELYEKTPFOOQLDAFAEKQALLNEHGATQEQLN 515
QY 541 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLRSQLVKRRKQNELRLQGELOKA 600
DB 516 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLRSQLVKRRKQNELRLQGELOKA 575
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631
DB 576 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 606

RESULT 3
US-08-402-217A-3
; Sequence 3, Application US/08402217A
; Patent No. 5587301
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
```



```
Db 301 LEKITDLKNQLOQDEDFRQLEEKGRKTAENVMTELTMEINKWRLLYEELYEKTPTF 360
QY 515 QOOLDAFAEAKQALLNEHGATQEQNLKIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKS 574
Db 361 QOOLDAFAEAKQALLNEHGATQEQNLKIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKS 420
QY 575 EVSKLSQLVKKRQKQNELRQLOGLDQALGIRHFDPSKAFCHASKENFTPLKEGNPNCC 631
Db 421 EVSKLSQLVKKRQKQNELRQLOGLDQALGIRHFDPSKAFCHASKENFTPLKEGNPNCC 477

RESULT 5
US-08-995-654-3
; Sequence 3, Application US/08995654
; Patent No. 6025138
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig
; APPLICANT: Seilhamer, Jeffrey
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/995,654
; FILING DATE: December 22, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/700,178
; FILING DATE: August 20, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/402,217
; FILING DATE: March 10, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0028-2 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-995-654-3

Query Match 75.8%; Score 2354; DB 3; Length 477;
Best Local Similarity 100.0%; Pred. No. 4e-135;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 MOILTERLALEREQVEKLOQKELQSQLSLQOQKELSAKELQSQLSQFQEMTSEKNVFKEE 214
Db 1 MOILTERLALEREQVEKLOQKELQSQLSLQOQKELSAKELQSQLSQFQEMTSEKNVFKEE 60
QY 215 LKLAELDAVQOKEEQSERLVKQLEERKSTAEOLTRLDNLLEKEVELEKHHIAHAQA 274
Db 61 LKLAELDAVQOKEEQSERLVKQLEERKSTAEOLTRLDNLLEKEVELEKHHIAHAQA 120
```

```
QY 275 ILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDV 334
Db 121 ILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDV 180
QY 335 TAOLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESYKSSTLKEIEDL 394
Db 181 TAOLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESYKSSTLKEIEDL 240
QY 395 KLENLTLOEKVAMAEKSVEDVQOQILTAESTNQEXYARMVQDQNLQRSTLKEEBEIKETTSF 454
Db 241 KLENLTLOEKVAMAEKSVEDVQOQILTAESTNQEXYARMVQDQNLQRSTLKEEBEIKETTSF 300
QY 455 LEKITDLKNQLOQDEDFRQLEEKGRKTAENVMTELTMEINKWRLLYEELYEKTPTF 514
Db 301 LEKITDLKNQLOQDEDFRQLEEKGRKTAENVMTELTMEINKWRLLYEELYEKTPTF 360
QY 515 QOOLDAFAEAKQALLNEHGATQEQNLKIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKS 574
Db 361 QOOLDAFAEAKQALLNEHGATQEQNLKIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKS 420
QY 575 EVSKLSQLVKKRQKQNELRQLOGLDQALGIRHFDPSKAFCHASKENFTPLKEGNPNCC 631
Db 421 EVSKLSQLVKKRQKQNELRQLOGLDQALGIRHFDPSKAFCHASKENFTPLKEGNPNCC 477

RESULT 6
US-08-402-217A-2
; Sequence 2, Application US/08402217A
; Patent No. 5587301
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,217A
; FILING DATE: 10-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF-002808
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-402-217A-2

Query Match 38.8%; Score 1205; DB 1; Length 351;
Best Local Similarity 59.6%; Pred. No. 6.9e-66;
Matches 258; Conservative 29; Mismatches 62; Indels 84; Gaps 1;

QY 155 MOILTERLALEREQVEKLOQKELQSQLSLQOQKELSAKELQSQLSQFQEMTSEKNVFKEE 214
||| :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 1 MNLKQKFILEQOEKREKLOQKELQDSLQKQKELSSHLQKLCFSQOEEMAKENKLFEE 60  
QY 215 LKLAELDAVQKQEEQSERLVKQLEERKSTAEQOLTRLDNLLREKEVELEKHIHAHAQA 274  
Db 61 LKOTDELQKQKQEEQSERLVKQLEERKSTAEQOLTRLDNLLREKEVELEKHIHAHAQA 120  
QY 275 ILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV 334  
Db 121 TLLLEKYSVMQSLDVT----- 139  
QY 335 TAQLESEQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV 394  
Db 140 -----AQFESYKALTASIEDL 156  
QY 395 KLENLTLOEKVAMARKSVEDVQOQILTAESTNOEYARMVQDQONRSTLKEEIKETTSF 454  
Db 157 KLENSLQEKVAKAGNAEDVQHQILATESSNOEYARMVQDQONRSTLKEEIKETTSF 216  
QY 455 LEKITDLNQLRQODEDFRKQLEEKGRKTAENYMTLTMEINKWRLLYEELYEYKTKPF 514  
Db 217 LQKITDLNQLRQODEDFRKQLEEKGRKTAENYMTLTMEINKWRLLYEELYEYKTKPF 276  
QY 515 QOQLDAFAEAKQALLNEHGAQEOQLNKIRDSYAQLLGHONLKQKIKHVVYKLDENSOLKS 574  
Db 277 QOQLDAFEVEKQALLNEHGAQEOQLNKIRDSYAQLLGHONLKQKIKHVVYKLDENSOLKS 336  
QY 575 EVSKLRSOLVKRK 587  
Db 337 EVSKLRCQLAKKK 349

RESULT 7

US-08-700-178-2  
; Sequence 2, Application US/08700178  
; Patent No. 5783669 5700912  
; Patent No. 5783669 5700912  
; GENERAL INFORMATION:  
; APPLICANT: Hawkins, Phillip R.  
; APPLICANT: Wilde, Craig G.  
; APPLICANT: Seilhamer, Jeffrey J.  
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN  
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/700,178  
; FILING DATE: August 20, 1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/402,217  
; FILING DATE: March 10, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0028-1 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 351 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-700-178-2  
Query Match 38.8%; Score 1205; DB 1; Length 351;  
Best Local Similarity 59.6%; Pred. No. 6.9e-66;  
Matches 258; Conservative 29; Mismatches 62; Indels 84; Gaps 1;  
QY 155 MOILTERALRQEVKEKLOQKELQDSLQKQKELSSHLQKLCFSQOEEMAKENKLFEE 214  
Db 1 MNLKQKFILEQOEKREKLOQKELQDSLQKQKELSSHLQKLCFSQOEEMAKENKLFEE 60  
QY 215 LKLAELDAVQKQEEQSERLVKQLEERKSTAEQOLTRLDNLLREKEVELEKHIHAHAQA 274  
Db 61 LKOTDELQKQKQEEQSERLVKQLEERKSTAEQOLTRLDNLLREKEVELEKHIHAHAQA 120  
QY 275 ILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV 334  
Db 121 TLLLEKYSVMQSLDVT----- 139  
QY 335 TAQLESEQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDV 394  
Db 140 -----AQFESYKALTASIEDL 156  
QY 395 KLENLTLOEKVAMARKSVEDVQOQILTAESTNOEYARMVQDQONRSTLKEEIKETTSF 454  
Db 157 KLENSLQEKVAKAGNAEDVQHQILATESSNOEYARMVQDQONRSTLKEEIKETTSF 216  
QY 455 LEKITDLNQLRQODEDFRKQLEEKGRKTAENYMTLTMEINKWRLLYEELYEYKTKPF 514  
Db 217 LQKITDLNQLRQODEDFRKQLEEKGRKTAENYMTLTMEINKWRLLYEELYEYKTKPF 276  
QY 515 QOQLDAFAEAKQALLNEHGAQEOQLNKIRDSYAQLLGHONLKQKIKHVVYKLDENSOLKS 574  
Db 277 QOQLDAFEVEKQALLNEHGAQEOQLNKIRDSYAQLLGHONLKQKIKHVVYKLDENSOLKS 336  
QY 575 EVSKLRSOLVKRK 587  
Db 337 EVSKLRCQLAKKK 349

RESULT 8

US-08-995-654-2  
; Sequence 2, Application US/08995654  
; Patent No. 6025138  
; GENERAL INFORMATION:  
; APPLICANT: Hawkins, Phillip R.  
; APPLICANT: Wilde, Craig  
; APPLICANT: Seilhamer, Jeffrey  
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN  
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/995,654  
; FILING DATE: December 22, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/700,178  
; FILING DATE: August 20, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/402,217







SEQUENCE CHARACTERISTICS:  
 LENGTH: 1886 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-938-105-3

Query Match 12.5%; Score 387; DB 4; Length 1886;  
 Best Local Similarity 23.9%; Pred. No. 1.le-15;  
 Matches 182; Conservative 129; Mismatches 239; Indels 210; Gaps 30;

QY 9 MRLNRKRETKMRSMVKGQGMELKIQ-ATQKD-LTESGKIVQLGKLVSTIEKEKIDK 66  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 779 MLYFKIPKPLKSAETEKEMANMKKEFGVKDALEKSEARRKELEKVMVSLQEKNDLQL 838  
 QY 67 ETEKLELYQIEISCASDVQVEKQDVIAOLE-----EDLKEKREILSLKSLSEN 116  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 839 QVQASQNDLAERCDQLRNKI---QLEAKVKEMTERLEDEEMNAELTAKRKLEDE 895  
 QY 117 IP-FSKQIEDLVKQQLLETNRDNLVSKDRERAETLSAEMQILTERLALEROEYKLOOK 175  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 896 CSELKKDIDDLTLAKVEKEKHATEN---VKNLTEMAGLDIIIAKLTKEKALQEA 951  
 QY 176 ELQSLSLQOKE-----LSARLQOOLCSFQEBMTSEKNV----- 210  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 952 HQAALDQLQAEEDKVTNLTLSKVKLEQQVDLDGSLQEKVKYRMDLERAKRKLGLKLT 1011  
 QY 211 -----PKELKALAEFDAVQ--KPEOSERLVKOLEERKSTAEQLTRLDNL 257  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1012 QESIMDLNDKIQLEKKKKEFDISQNSKIEDQALQLOKKL---ENQARIEEL- 1067  
 QY 258 REKEVELEKHAHAQAIIAQAQYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQL 317  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1068 -EELEAERTAKRQVLE-----RSDLTRELEISERLEEAGGA---TSVQI 1110  
 QY 318 ESEQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVS 377  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1111 ENMKKRAEFQKMRD---LEEATLOHEATAAALRKHA--DSVAE-LGEQIDNLRQVK 1163  
 QY 378 AQLESYKSKSLKEIEDL-----KLENL--TLQE-----KVMAEKSVEDVQ 416  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1164 QKLEKESEFKLEDDVSHMEQIIKANKLEKVSRTLEDQANEYRVKLEAQRSLNDFT 1223  
 QY 417 QOILPAESTNOBYARWQDLQNRSLKKEEIKETS-----SFLEKITDLKNLRQO----- 468  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1224 TORAKLOTENGELARQLEE-----KEALIWLTRGLSYTQQMEDLKRQLEEBGKAKN 1276  
 QY 469 -----DED-FRQLEBKGRTAENKENVMTLWETWINKRWLLY-----EELY 508  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1277 ALAHALQASRHDCDILREQEYEEEMAKAELQRLVLSKANSEVAQWRKYETDAIQRTTELE 1336  
 QY 509 EKTQPFQOOL-DAPEA-----EKOALLNE-----HGATQEQLNKTRD 544  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1337 EAKKLAQRLQDAEBAVAENAKCSSLEKTKHRLQNEIETDLMDVVERNAALADKKQR 1396  
 QY 545 SYAQLLGHQNLK-----RKQNE---LRQGEIDKA 600  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1397 NFDKILAEWKQYBESQSELESSQKARSLSLETFKLKNAYEESLEHLETFKRENKNLQE 1456  
 QY 575 EYVKLRSQLVK-----RKQNE---LRQGEIDKA 600  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1457 EISDLTEQLGEGGKNVHLEKIRKOLEVEKLEQALQEA 1496

RESULT 13  
 US-09-572-191-2  
 ; Sequence 2, Application US/09572191  
 ; Patent No. 6355466  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; APPLICANT: Sakowicz, Roman  
 ; APPLICANT: Wood, Kenneth  
 ; TITLE OF INVENTION: No. 6355466el motor proteins and methods for

; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1017  
 ; CURRENT APPLICATION NUMBER: US/09/572,191  
 ; CURRENT FILING DATE: 2000-05-17  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-09-572-191-2

Query Match 12.1%; Score 375; DB 4; Length 1388;  
 Best Local Similarity 22.8%; Pred. No. 4.2e-15;  
 Matches 165; Conservative 142; Mismatches 230; Indels 188; Gaps 28;

QY 3 ALSLELMKLRNKRRETKMRSMVKGQGMELKIQATQKDTESGKIVQLGKLVSTIEKSKI 62  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 714 AISSELRTVQ-EQMSALQAKLDEEHKKNLQHQHVDKLEHSHSTQMOEL-----FSSERI 766  
 QY 63 DEKCETEKLL-----EYIOEISCASD---QVEKCKVDIAQLEEDLKEKDRITLSLQ 111  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 767 DWTQOQEBLLSQLNVLEKQLQETQTKNDFLAKSEVHDRLRVLHSDAKELSSVVKLEYSSFT 826  
 QY 112 SLEENITFSQIE-DLAVKCOL-----LETER-----DNL----- 140  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 827 NOEKE--FNKLSERHMHVOLQLDNLRLENKLLSKACLOQSDYNDLQELMKFEIDOLSRN 884  
 QY 141 VSKDRERAEETLSAEMQILTERLALEROEYKL-----QOKELOSQSL-----QOKE 188  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 885 LONFKKETLTKSLDNLNMLLEAEKERNKLSLQFEEDKENSKEILKLVAVNRQEKOR 944  
 QY 189 LSARLQOOLCSFQEBMTSEKNVFEELKALAEALDAVQOKEEQSER-----LVKOLEER 243  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 945 ETACEQOMAKVQK-----LEESLLATEKYVSSLEKSRSDKVVVADLMNQIOELR 995  
 QY 244 KSTAEQLTRLDNLLRE-KEVELEKHIA---AHAQAIIAQAQYNDTAQSL----- 289  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 996 SSVCEKTETIDTLQELKQKINDCKYNSALVDREESRVLLIKQEVDTLDLKETRLRLILSED 1055  
 QY 290 --RQVTAQ-LESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYN 346  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1056 IERDMLCEDLAHATEQLNMLTEASKKSHGLQSAQEEELTKKALIQELQHLNOKKEVE 1115  
 QY 347 DTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYN 386  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1116 QKKNYFNKMQLEHVMDAAEDQSPKTPPHFQTHLAKLLETQEOEIEDGRASKTSLEH 1175  
 QY 387 -----TLKEIEDLKLLENLTQLEKVAEKSVEDVQOQIITAE 425  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1176 LVTKLINEDREVKNABEILRMKEQLREMENTLRESQOLIEKNWLLQGLDDIKRQ---KENS 1232  
 QY 426 NOEYARVQDQLNRLSTLKEEIKE-----ITSSFLEKITDLKNQ-----LRQO 468  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1233 DQNH-----PDNQOLNKEQEEESIKERLAKSIVEEMLMKMAKDLVEVQSALYNKMECLRMT 1288  
 QY 469 DEFDRKOLEEKGKRTAEKENVMTELTEINKWRLLYELYEKTPPQOQLDAFAEKQAL 528  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1289 DEVERTQTLT-SKAFQEKQELRSKL-----EEMEYERERTSQEMENLRKQVBECL 1336  
 QY 529 LNEHGATQEQNLKIRDSYAQLLGHQNLKQIKHVVKLKDENSQKSEVSKLRSQ---LVK 585  
 ||| | : : : : : ||| | : : : : : ||| | : : : : : ||| | : : : : :  
 Db 1337 AEENG-----KLVGHQNLHOKIOYVVVRLKENVRLAAEETEKLAENVFUKE 1382  
 QY 586 RKQNE 590  
 ||| | : : : : :  
 Db 1383 KKRSE 1387

RESULT 14  
 US-09-723-262-2  
 ; Sequence 2, Application US/09723262  
 ; Patent No. 6379912

[illegible]

QY 529 LNEHGATQEQLNKIRDSYAQLLGHQNLKQKIRKHVVVKLDKDNSQLKSEVSKLRSQ----LVK 585  
Db 1337 AEENG-----KLVGHQNLHOKIQYVVVLKKNVRLAEETEKLRANVFLKE 1382

QY 586 RKQNE 590  
Db 1383 KKRSE 1387

Search completed: December 20, 2002, 15:14:24  
Job time : 21.0295 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:09:43 ; Search time 31.6431 Seconds  
(without alignments)  
4108.822 Million cell updates/sec

Title: US-09-685-010-48  
Perfect score: 3104  
Sequence: 1 MRALSLEMLKLNKRKRETKMR.....FCHASKENFTPLKEGNPNCC 631

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3048	98.2	836	11	Q8VDR2 mus musculus
2	2374.5	76.5	713	11	Q9WUF7 rat
3	2338.5	75.3	476	11	Q920Z6 mus sp. hya
4	616.5	19.9	240	6	Q8SPR5 bos taurus
5	437.5	14.1	1598	11	Q922D2 mus musculus
6	431.5	13.9	1790	3	Q07380 saccharomyc
7	428.5	13.8	1937	6	Q9TV62 sus scrofa
8	425	13.7	1940	6	Q9BE41 bos taurus
9	423	13.6	1939	6	Q9TV63 sus scrofa
10	418	13.5	1944	13	Q9DGM5 gallus gall
11	417	13.4	1938	6	Q9BE40 bos taurus
12	415	13.4	1410	4	Q14221 homo sapien
13	414	13.3	1411	4	Q15075
14	413	13.3	1939	13	Q9PTY2 gallus gall
15	411	13.2	1939	6	Q9TV61 sus scrofa
16	408.5	13.2	1690	5	O44929 drosophila

17	405	13.0	1941	13	Q9DGM4	Q9dgm4 gallus gall
18	402.5	13.0	1935	13	Q90ZE5	Q90ze5 brachydanio
19	402	13.0	1938	13	Q9IBD7	Q9ibd7 seriola dum
20	401.5	12.9	1933	13	Q90337	Q90337 cyprinus ca
21	399.5	12.9	1950	5	Q26080	Q26080 placopecten
22	398.5	12.8	1941	5	Q26079	Q26079 placopecten
23	397	12.8	1368	13	Q9PTD7	Q9ptyd7 xenopus lae
24	396.5	12.8	1690	5	Q9VJE5	Q9vje5 drosophila
25	395.5	12.7	1937	13	Q9IBD4	Q9ibd4 gallus gall
26	395	12.7	1932	13	Q98TQ4	Q98td4 notothenia
27	395	12.7	1938	6	Q9GJF9	Q9gjf9 oryctolagus
28	393	12.7	1463	5	Q9GY20	Q9gyz0 strongyloce
29	391	12.6	2954	13	Q42263	Q42263 xenopus lae
30	390.5	12.6	1092	13	Q90338	Q90338 cyprinus ca
31	390.5	12.6	1931	13	Q42352	Q42352 cyprinus ca
32	390	12.6	1738	5	O76329	O76329 dictyosteli
33	389	12.5	1929	13	Q98TQ6	Q98td6 notothenia
34	388.5	12.5	1935	4	Q9H1D5	Q9h1d5 homo sapien
35	388	12.5	1958	5	O96062	O96062 dugesia jap
36	387.5	12.5	1583	4	O15045	O15045 homo sapien
37	387	12.5	1960	11	O8VDD5	O8vdd5 mus musculu
38	386.5	12.5	1935	6	Q9BE39	Q9be39 bos taurus
39	386	12.4	1930	13	Q9DGD5	Q9dgd5 pennahia ar
40	384.5	12.4	2139	5	Q07569	Q07569 entamoeba h
41	384	12.4	1936	13	Q90YF6	Q90yfe paracirrhith
42	384	12.4	1979	5	O96133	O96133 plasmodium
43	383.5	12.4	1931	13	Q91OC5	Q91oc5 gallus gall
44	383.5	12.4	1935	6	Q9GKRL	Q9gkrl sus scrofa
45	381	12.3	1935	11	Q91283	Q91283 mus musculu

ALIGNMENTS

RESULT 1

Q8VDR2 ID Q8VDR2 PRELIMINARY; PRT; 836 AA.  
AC O8VDR2;  
DT 01-MAR-2002 (TREMblrel. 20, Created)  
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE Similar to hyaluronan mediated motility receptor (RHAMM).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC021427; AAH21427.1; -  
KW Receptor.  
SQ SEQUENCE 836 AA; 96670 MW; B9BDDA22EBECA652 CRC64;

Query Match		98.2%;	Score 3048;	DB 11;	Length 836;
Best Local Similarity		93.0%;	Pred. No. 2.8e-101;		
Matches 626;		Conservative 2;	Mismatches 3;	Indels 42;	Gaps 1;
Qy	1	MRALSLEMLKLNKRKRETKMRSMVQEGMELKQATQKDLTSGKIVQLEKGLVSI	IEKE 60		
Db	164	MRALSLEMLKLNKRKRETKMRSMVQEGMELKQATQKDLTSGKIVQLEKGLVSI	IEKE 223		
Qy	61	KIDKCEFEKLELEYTOEISCSASQVQKCKVDIAQLEEDLKEKDREILSLKQSL	ENITFS 120		
Db	224	KIDKCEFEKLELEYTOEISCSASQVQKCKVDIAQLEEDLKEKDREILSLKQSL	ENITFS 283		
Qy	121	KQIEDLTVCQQLLETERNLVSKDRERATLSAEQMLTERLALEREQYEKIQKEL	QSQ 180		
Db	284	KQIEDLTVCQQLLETERNLVSKDRERATLSAEQMLTERLALEREQYEKIQKEL	QSQ 343		
Qy	181	SLLQKEKELSLARLQQLCSFQEMTSEKNVFKELKALAEALDAVQKKEQSER	LVKQLE 240		
Db	344	SLLQKEKELSLARLQQLCSFQEMTSEKNVFKELKALAEALDAVQKKEQSER	LVKQLE 403		

```
QY 241 ERKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIA----- 278
Db 404 EETKSTAEQLKRLDNLREKEVELEKHIAHAQAAILIAEQKYNDAQSLRDVTAQLESQ 463
QY 279 -----QKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQ 318
Db 464 EKYNDTAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQ 523
QY 319 SEQEKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQ 378
Db 524 SEQKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQ 583
QY 379 QLESYKSSLTKEIDLKLENLTLOEKVMAEKSVEDVQQOILTAESTNOEYARMVQDLQN 438
Db 584 QLESYKSSLTKEIDLKLENLTLOEKVMAEKSVEDVQQOILTAESTNOEYARMVQDLQN 643
QY 439 RSTLKEEETKEITSSPLEKITDLKNQROODEDFRKQLEEKRAEKENVMTLMEIN 498
Db 644 RSTLKEEETKEITSSPLEKITDLKNQROODEDFRKQLEEKRAEKENVMTLMEIN 703
QY 499 KRWLLYEELYEYKTPFQOOLDAFAEKQALLNEHGATQEQLNKIRDSYAQLLGHONLKOK 558
Db 704 KRWLLYEELYEYKTPFQOOLDAFAEKQALLNEHGATQEQLNKIRDSYAQLLGHONLKOK 763
QY 559 IKHVYKLDENSQKSEVSKLSQVKKRQNELRLQGLDQKALGIRHFDPSKAFCHASKE 618
Db 764 IKHVYKLDENSQKSEVSKLSQVKKRQNELRLQGLDQKALGIRHFDPSKAFCHASKE 823
QY 619 NETPLKEGPNCC 631
Db 824 NETPLKEGPNCC 836

RESULT 2.
Q9WUF7 PRELIMINARY; PRT; 713 AA.
AC Q9WUF7;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Hyaluronan receptor RHAMM.
GN RHAMM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY;
RA Li X., Lynn B., Nagy J.I., Cattini P.A.;
RL "RHAMM cDNA from rat brain.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PHOCHROMOCYTOMA;
RA Lynn B.D., Li X., Cattini P.A., Nagy J.I.;
RT "Sequence, protein expression and erk association of the hyaladherin
RHAMM in PC12 cells.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF133037; AAD24473.1; -
DR EMBL; AF336825; AAK21904.1; -
KW Receptor.
SQ SEQUENCE 713 AA; 82395 MW; 2FE310D5759C6CB3 CRC64;

Query Match 76.5%; Score 2374.5; DB 11; Length 713;
Best Local Similarity 77.4%; Pred. NO. 1.9e-77;
Matches 491; Conservative 28; Mismatches 28; Indels 87; Gaps 2;

QY 1 MRALSLEMLKLNKRKTRKMSMMVKQEGMELKQATQKDLTFSKGIYQLEGKLSIEKE 60
Db 164 MTLSEMLKLNKRKTRKMSMMVKQEGMELKQATQKDLTFSKGIYQLEGKLSIEKE 223
```

```
QY 61 KIDECETEKLEYTOEISCASQDQVEKCKVDIAQLEEDLKEKDREILSKQSEENITPS 120
Db 224 KIDECETEKLEYTOEISCASQDQVEKCKVDIAQLEEDLKEKDREILSKQSEENITPS 283
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAFSLSAEQWILTERLALEROEYKIQQLEQSO 180
Db 284 KOIEDLTVKCOLLETERDNLVSKDRERAFSLSAEQWILTERLALEROEYKIQQLEQSO 343
QY 181 SLLOQEKELSARLOQLCSFQEBMTSEKNVFKKEELKLAELDAVQOQKEQSERLVKQLE 240
Db 344 SLLOQEKELSARLOQLCSFQEBMTSEKNVFKKEELKLAELDAVQOQKEQSERLVKQLE 403
QY 241 ERKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIAEQKYNDAQSLRDVTAQLESQ 300
Db 404 EETKSTAEQLRRLDNLREKEVELEKHIAHAQAATVIAQEKYSDTAQLRDVT----- 456
QY 301 EKYNDTAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQ 360
Db 457 -----QKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQKYNDAQSLRDVTAQLESQ 456
QY 361 SVOEKYNDAQSLRDVSAQLESYKSSLTKEIEDLKLENLTLOEKVMAEKSVEDVQQOIL 420
Db 457 -----AQLESYKSSLTKEIEDLKLENLTLOEKVMAEKSVEDVQQOIL 499
QY 421 TAESTNOEYARMVQDLNRLSTLKEEIEITSSFLKEITDLKNQROODEDFRKQLEEK 480
Db 500 TAESTNOEYARMVQDLNRLSTLKEEIEITSSFLKEITDLKNQROODEDFRKQLEEK 559
QY 481 KRTAEKENVMTLMEINRWLLYEELYEYKTPFQOOLDAFAEKQALLNEHGATQEQLN 540
Db 560 AKMTEKETAVTETLMEINRWLLYEELYEYKTPFQOOLDAFAEKQALLNEHGATQEQLN 619
QY 541 KIRDSYAQLLGHONLKQKIKHVYKLDENSQKSEVSKLSQVKKRQNELRLQGLDQKAL 600
Db 620 KIRDSYAQLLGHONLKQKIKHVYKLDENSQKSEVSKLSQVKKRQNELRLQGLDQKAL 679
QY 601 LGIRHFDPSKAFCHASKENF---TPLKEGPNCC 631
Db 680 LGIRHFDPSKAFCHASKENF---TPLKEGPNCC 713

RESULT 3
Q92026 PRELIMINARY; PRT; 476 AA.
AC Q92026;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Hyaluronan receptor.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92299690; Pubmed=1376732;
RA Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,
RA Cripps V., Austen L., Nance D.M., Turley E.A.;
RT "Molecular cloning of a novel hyaluronan receptor that mediates tumor
cell motility.";
RL J. Cell Biol. 118:753-753(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=92348516; Pubmed=1639856;
RA Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,
RA Cripps V., Austen L., Nance D.M., Turley E.A.;
RT "Correction. Molecular cloning of a novel hyaluronan receptor that
mediates tumor cell motility.";
RL J. Cell Biol. 118:753-753(1992).
DR EMBL; S41029; AAA09809.1; -
KW Receptor.
SQ SEQUENCE 476 AA; 55486 MW; 62DCA82045EE0CB2 CRC64;
```



```

QY 440 STLKEE-----ETKEITSSFELEKTTDLKNQLRQQ----- 468
Df 1285 GRLOTGEFGRDLDKEALYSQLSRGQAQTQOITEELKROLEEVEVKAKNALAHALQSRR 1344
QY 469 -DED-FRKOLEEKGRTAENNVNMTLMEINKWRLLY-----BELYEKTFFPOOL 518
Df 1345 HCDCLLRQYEBOESKAELQRALSKANSEVAQWRTKYETAIO RTEEEAAKKLAORL 1404
QY 519 DA-----FAEKOALLNE-----HGATOBQLN----- 540
Df 1405 QAABEHVAVNAKCASLEKTKORLQNEVEDIMLDVERTNAACAALDKQRNFDDKILAEWK 1464
QY 541 -KIRDSYAQL-----LGHQNLKQK-----IKHVYKLKDENSEQLSKSEVSLRSOL- 583
Df 1465 QKYETHAELEASOKEARSLGTELFMKMNAVEESLDQLETCLKRENKNIQQEISDLTEQIA 1524
QY 584 -----VKR--KQNELRLQGLDKA 600
Df 1525 EGGKRHIELEKIKKQVEQECLEQAALAEA 1554

RESULT 6
Q07380 PRELIMINARY; PRT: 1790 AA.
ID AC Q07380 P89892;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 206.5 kDa protein YDL058w.
GN USOI OR YDL058w.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_Taxid=4932;
RN [1]
RP RPA SEQUENCE FROM N.A.
RA Bloembergen H., Brandt P.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL: Z74106; CAAG8621.1; -.
DR EMBL: Z74105; CAAG8620.1; -.
DR SGD: S0002216; USOI.
DR InterPro: IPR002017; Spectrin.
SQ SEQUENCE 1790 AA; 206450 MW; 90062544F5A52EE CRC64;

Query Match 13.9%; Score 431.5; DB 3; Length 1790;
Best Local Similarity 24.2%; Pred. No. 3 Be-08;
Matches 179; Conservative 149; Mismatches 236; Indels 177; Gaps 33;

QY 10 KLRNKRETKMRSMVMYQEGMELKLOAQTKDLTESKGIVQLLEGKLVSTIEKEKIDEKCETE 69
Df 798 ELKNVRDS-LDEMTOLRDVLTCKENQTAALLEYKSTHKQEDSIIKTLEK----- 846
QY 70 KLLEYIQEISCASDOVEKCKVDIAQLLEDL-----KEKD-----REILSUQ 111
Df 847 GLETLISQKKAEDGINMGMDLEFALSREMGAVEENCNKLQEKDKSNVNHQKETSLSKE 906
QY 112 SLEENIY----FSKQIEDLTGVKCOLLETER-----DNLYSKDRERAET 150
Df 907 DIAAKITEIKAINLENLEMKTQCNNLSKERBHISKELVEYKSRFQSHDNVLAKUTEKLS 966
QY 151 LS---AEMQILTERL--ALERQEYE-KLQKELQSQ-SLLQOEKELSARLOQLCSFOEE 203
Df 967 LANNYKDMQAEENSLIKAVEESKNESSTQLSNLQNKIDSMSQKE-----NQIE 1016
QY 204 MTS-EKNVFVEELKLAELAELDAVOOK-----EEQSERLIVQLEERKSTABQL 250
Df 1017 RGSTIEKNI--EQLKTIISDLTEQTKEEIISKSDSKDEYESQISLLKEKLETATTANDENV 1074
QY 251 TRDNLAREVELEKHIAHAQAAILIAQEKYNNTAQSRLRDVTQALESVQE-----KY 303
Df 1075 NKISELTKTRE-EUEAEUAYKNUKNEULETSETSEKALKREVENEHLKEKTELKEEA 1133
QY 304 NDTAQSRLRDVTQALESQEKYNNTAQSRLRDVTQALESQEKRYNDTAQSLRD-VTA--QL 359

```

Db	1134	TETKQOLNSLRANLESLKEHEHDLAAQKKYEEQIANKEROYNEEISQLNDEITSTQOEN	1193
Qy	360	ESVQEKYND--TAQSLRDVSAQLESYKSTLK-----EIEDLKLNTL-----	400
Db	1194	ESIKKNDLEGEVGVKAWKSTSEQSNLKSEIDALNLQIKELKKNETNEASLESISKV	1253
Qy	401	-----LQEKVMAEKSVEDVQOQIILTAESTNQEYARMVQDQLNKSTLKKEEIKREIT	451
Db	1254	ESETVKIKELQDCENCFEKEVSELEDKLSAKSDKSKYL-----ELQEKSEKIKEELDAKT	1309
Qy	452	SSP-----LEKITDLKLNLRQODEDFRKLQLEBKGRKTAERKENYMTLTHEINKRWLLYBELY	508
Db	1310	TELKIQLEKITTNL-----SKAKESESE--LSRLKKTSSPERKNAEEQL	1351
Qy	509	EKTK-PPQOOLDAPEAKQALLNEHCAT-----QQLANKIRDSYAQLLGHONLKQKIKHV	562
Db	1352	EKLKNETQTKNAQAFERK-LNKGSSITTYQSEKINTLEDELRILQONENELAK----	1406
Qy	563	VKLKDNSOLKSVSRSLRSQLVKRRQNELR-LQGSLL-----DKALGI-----RH	605
Db	1407	EIDNTRSELE-KVSLNDELLEEKQNTIKSLQDEILSYKDKITRNDEKLLSIERNKRD	1464
Qy	606	FDPSPKAFCHASKENFTPLKEG	626
Db	1465	LESLEKQLRAAQESKAKVEEG	1485

RESULT 7

Q9TV62

ID

Q9TV62

PRELIMINARY;

PRT; 1937 AA.

AC

Q9TV62;

DT

01-MAY-2000 (TReMBLrel. 13, Created)

DT

01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT

01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE

Myosin heavy chain 2b.

OS

Sus scrofa (Pig).

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX

NCBI\_TaxID=9823;

RN

[1]

SEQUENCE FROM N.A.

STRAIN=LANDRACE; TISSUE=SKLETAL MUSCLE;

RC

Chikuni K., Tanabe R., Muroya S., Nakajima I.;

RT

"Differences in molecular structure among the porcine myosin heavy

RT

chain-2a, -2x, and -2b isoforms.;

RL

Meat Sci. 57:311-317(2001).

ENBL

AB025261; BAA82145.1; -.

DR

HSSP; PI3538; 2MYS.

DR

InterPro; IPR000048; IQ\_region.

DR

InterPro; IPR001609; myosin\_head.

DR

InterPro; IPR004009; Myosin\_N.

DR

InterPro; IPR002928; Myosin\_tail.

DR

Pfam; PF00612; IQ; 2.

DR

Pfam; PF00063; myosin\_head; 1.

DR

Pfam; PF02736; Myosin\_N; 1.

DR

Pfam; PF01576; Myosin\_tail; 1.

DR

PRINTS; PR00193; MYOSINHEAVY.

DR

ProDom; PD000355; myosin\_head; 1.

DR

SMART; SM00015; IQ; 1.

DR

SMART; SM00242; MYSC; 1.

DR

PROSITE; PS50096; IQ; 1.

SQ

SEQUENCE 1937 AA; 232335 MW; BBC114C6824E0426 CRC64;

Query Match

13.8%; Score 428.5; DB 6; Length 1937;

Best Local Similarity

22.5%; Pred. No. 5,2e-08;

Matches 170; Conservative 154; Mismatches 248; Indels 185; Gaps

Qy	9	MKLNRRETKMRSMVVKOE-GMELKLAQTKDLTESKGIQVLESGKLVSTKEKIDEKC	66
Db	832	MKLIYFKIPLLSAETEKEMANNEFEKTDLEAKSKRKELEKRWALMOKNDQLQ	891
Qy	67	ETEKILEYIQEISCDASDVVEKCKV-----DIAQLEDLKEKREILTSIKQSLEENIT-	118



```
Db 892 QVQAEADGLADAEERCDQLIKTQLEAKIKVENTRAEDEEBEINAEITAKKRKLEDEGSE 951
Qy 119 FSKQIEDLVTKCOLLETERDNLVSKDRERAETLSAEMOILTERLALEREQYEKLOQKELQ 178
Db 952 LKKDIDDLLELTLAKVEKEKHATENK-----VKNLTEEMAGLDENIAKLTKEKALQEAHQ 1007
Qy 179 SOSLIQOEKE-----LSARLOQOLCSFOEEMTSEKNV-----FRELKAL-A 220
Db 1008 TLDDLOAEEDKVNTLTAKTKLEQQVDDLEGSLEQEKLRMDLERAKRKLEGLKLAQES 1067
Qy 221 ELDAVOQEESERLVKQ-----LE-----ERKSTAEQOLTRDNLRLREKEVELEKHIAAH 271
Db 1068 TMDIENDKQOQDLKLEKKEFEFMSNLQSKIEDEQALAMOLKIKELQARTAELEEEIEAE 1127
Qy 272 AQAILIAOEKYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESQEKYNDVTAQSL 331
Db 1128 RASRAKAEKQSDLSRELEIEISERLE-----EAGGATSAQIEMNKKREAEFKMR 1177
Qy 332 RDVTAQLESQEKYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESYKSSTLKEI 391
Db 1178 RD-----LEEATLQHEATAAALRKKHA--DSVAE-LGEQIDNLRVQKQLEKEKSELKWEI 1230
Qy 392 EDLKLNLTLQEKVMAEKSVEDVQOQILTAESTNQEYARVWQDLOKRLSTLKEEBIKIT 451
Db 1231 DDLASNMETVSKAGNLEKMCRTLEDQLSYKTRKEEHQRLINELSAQKARLQTESGFS 1290
Qy 452 -----SSFEKITDKNLQROQ-----DED-FRKOLE 477
Db 1291 QOLDEKEALVSQLSRGKQAFQOIEELARQLEETKAKSALAHAVOSSRHDCDLLREQYE 1350
Qy 478 EKGKRTAEKENVMTELTMEINKRWLLY-----EELYEKTQPFQOQLD----- 519
Db 1351 EEQEAKEALQRAKANSANSEVAQWRTKYETDAIQTREELEEAQKLAQRLQDAEHEVAVN 1410
Qy 520 -----AFEAKEQALLNE-----HGATQEOQLN 540
Db 1411 AKCASLEKTRQRLQNEVEDLMDVRSNAACAALDKKQNFDFKILAENKHKRYEETQAELE 1470
Qy 541 -----KIRDSYAQLLG-----HONLKOKI-----KHVVKLKD 567
Db 1471 ASQKESRSLSTELFKVKNAYEESLDQLETLRKKNLQOEISDLTQEAEGGKHIHELEK 1530
Qy 568 ENSQLKSEVSKLRSQVKKQNELRQELGDKALGIR 604
Db 1531 VKKQIEQEKSELQAL-----EEAEASLEHEEGKILRIQ 1564

RESULT 8
Q9BE41 ID Q9BE41 PRELIMINARY; PRT; 1940 AA.
AC Q9BE41;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Myosin heavy chain 2a.
GN MYHC-2a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HOLSTEIN; TISSUE-SKELETAL MUSCLE;
RA Chikuni K., Muroya S., Nakajima I.;
RT "Sequencing of the bovine myosin heavy chain isoforms."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB059398; BAB40920.1; -.
DR HSSP; P13538; 2MYS
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
```

```
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00663; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 1.
SQ SEQUENCE 1940 AA; 223318 MW; EE2642E1E29FDDC2 CRC64;

Query Match 13.7%; Score 425; DB 6; Length 1940;
Best Local Similarity 23.1%; Pred. No. 6.9e-08;
Matches 169; Conservative 143; Mismatches 25; Indels 162; Gaps 23;

Qy 4 LSLELMKLRNKRRETKMRSMVMYKQEGMELKLQATQDQLTESKGIYQLEGKLVISIEKEID 63
Db 837 LFFRIKPLLLKSAETEKEMATKKE-----FOKTKDELAKSEAKRKELEKVMYTLLEKND 891
Qy 64 EKCETEKLLYIQETSCASDOVEKCV-----DIAOLEEDLKEKDRILSLKQSLLEN 116
Db 892 LOLQVSAEAGLADAEERCDQLIKTQLEAKIKVENTRAEDEEBEINAEITAKKRKLEDE 951
Qy 117 IT-FSKQIEDLVTKCOLLETERDNLVSKDRERAETLSAEMOILTERLALEREQYEKLOQ 175
Db 952 CSELKDDIDDLLETLAKVEKEKHATENK-----VKNLTEEMAGLDETIAKLTKEKALQEA 1007
Qy 176 ELQOSLILQOEKE-----LSARLOQOLCSFOEEMTSEKNV-----FRELKLA 218
Db 1008 HQQTLDLLOAEEDKVNTLTAKTKLEQQVDDLEGSLEQEKLRMDLERAKRKLEGLDKLA 1067
Qy 219 LAELDVAQOKEESERLVKQLE-----EERKSTAEQOLTRDNLRLREKEVELEKHI 268
Db 1068 QESINDIENEKQDLDEKLLKKEFEISNLQSKIEDQALGILQKKIKELQARIEELEEEI 1127
Qy 269 AAHAQAIIIAOEKYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESQEKYNDVTA 328
Db 1128 EAERASRAKAEKQSDLSRELEIEISERLE-----EAGGATSAQIEMNKKREAEFQ 1177
Qy 329 QSLRDVTAQLESQEKYNDVTAQSLRDVTAQLESVOEKYNDVTAQSLRDVTAQLESYKSSTL 388
Db 1178 KMRRD-----LEEATLQHEATAAALRKKHA--DSVAE-LGEQIDNLRVQKQLEKEKSEM 1230
Qy 389 KEIEDLKLNLTLQEKVMAEKSVEDVQOQILTAESTNQEYARVWQDLOKRLSTLKEEBIK 445
Db 1231 MEIDDLASNVETISKAGNLEKMCRTLEDQVNLKSEKEEQOQLINDLTQGRQLQTESG 1290
Qy 446 -----EIKEITSSFEKITDKNLQROQ-----DED-FRK 474
Db 1291 EFSRQLEKEALVSQLSRGKQAFQOIEELKQLEEEIKAKNALAHGLQSAHRDCDLLRE 1350
Qy 475 QLEEKGRKTAENVMTELTMEINKRWLLY-----EELYEKTQPFQOQLD----- 520
Db 1351 QYEEQESKAEALQRAKANSANSEVAQWRTKYETDAIQTREELEEAQKLAQRLQDAEHEV 1410
Qy 521 -----FEAEKQALLNE-----HGATQEOQLN-----KIRDSYA 547
Db 1411 AVNAKASLEKTRQRLQNEVEDLMDVETNNAACAALDKKQNFDFKILAENKHKRYEETHA 1470
Qy 548 QI-----IGHONLKQK-----IKHVVKLKDENSEKSKSEVSKLRSQVKK--RKQNE 590
Db 1471 ELEAAQKEARSIGTFLFKMKNAYEESLDQLETLRKKNLQOEISDLTQEAEGGKRMHE 1530
Qy 591 L---RQELDKA 600
Db 1531 LEKIKKQVEQES 1543

RESULT 9
Q9TV63 ID Q9TV63 PRELIMINARY; PRT; 1939 AA.
AC Q9TV63;
```

```
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Myosin heavy chain 2a.
OS Mus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LANDRACE; TISSUE=SKELETAL MUSCLE;
RA Chikuni K., Tanabe R., Muroya S., Nakajima I.;
RT "Differences in molecular structure among the porcine myosin heavy
RL chain-2a, -2x, and -2b isoforms.";
DR Meat Sci. 57:311-317(2001).
DR EMBL; AB025260; BAA82144.1; -.
DR HSSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ_region..
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1..
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
DR SQ SEQUENCE 1939 AA; 223149 MW; B520182094179343 CRC64;

Query Match
Best Local Similarity 13.6%; Score 423; DB 6; Length 1939;
Matches 170; Conservative 147; Mismatches 254; Indels 168; Gaps 23;

QY 9 MKLRNKRFTKMSVMVKOE--GMELKQATQKDLTQESKGIQVLEGKLVSEIKSDEK 66
DB 834 MKLFKIKPLKLSAESEKEMANMKEEFKTKELAKSAKRKELEKMMVTLKKNLQL 893
QY 67 ETEKLELYIQEISCASDQVEKCV-----DIALEEDLKEKREILSLKOSLEENIT- 118
DB 894 QVQAEADALAEERCQDLIKTKIQLEAKIKEVTERAEDEEINAEIATAKRRKLEDCSE 953
QY 119 FSKQIEDTVKQCLLETDRNLVSKDRERAETLSAEMQILTERALEROEYKQKQELQ 178
DB 954 LKKDIDDLTLAKYKEKHATENK----VKNLTEEMAGLDETIAKTKKALQEAHQ 1009
QY 179 SOSLQOQKE-----LSARLQOQLCSFOEEMTSEKNV-----FKELKLALAE 221
DB 1010 TLDDLQAEEDKYNVTUTKAKTKLEQVDDLEGSLEQEKLRMDLERAKRKLGLDLKLAQES 1069
QY 222 LDVQOQKEQSRVLKQLE-----EERKSTAEQLTRLDNLLREKEVELEKHIAH 271
DB 1070 IMDIENEKQOLDEKKKEFEFETSNLSQKTEDEQALAIQKKIKELQARIELEEEIAR 1129
QY 272 AQAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSL 331
DB 1130 RASRAKAEQRSDLSRELEIEISLERLE-----EAGGATSAQIEMNKKRAEFQKMR 1179
QY 332 RDVTAQLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESYKSTLKEI 391
DB 1180 RD-----LEEATLQHEATAALRKHA--DSVAE-LCEQIDNLRQVKQLEKEKSEMKMEI 1232
QY 392 EDLKLLENLTQEKVMAEKSVEDVQOQILTAESTNQEYARMYQDL-QNRSTLKEE---- 445
DB 1233 DLASNMETVSKAGNLKRCMTLEDQLSLKSEEEQORLINDTAQGRQLQTSSEGFES 1292
QY 446 -----BIKETTSFLEKIKITDLKNQLRQ-----DED-FRKOLE 477
DB 1293 RQLEDEKALVSLSRGQAYTQIQIEELKRLQEEIIEKAKNALAHALQSSRHCDLLREQYE 1352
QY 478 EKGKRTAEKENVMTELTMWINKWRLLY-----EELYETKTPFQQQLDA----- 520
```

```
DB 1353 EQESKAEQLRALSKANTEVAQWRTKYETDAIQTREELEAKKLAQRLQAAEEHVEAVN 1412
QY 521 -----FEAEKQALLNE-----HGATQEQQLN-----KIRDSVAQL- 549
DB 1413 AKCASLETKTKORLQNEVEDMLDVERTNACAALDKORNFDKIILAEWKQKYEETHAELE 1472
QY 550 -----LGHONLKOK-----IKHVVKLKQDENSQKSEVSKLSRQL----- 583
DB 1473 ASQKEARSLSGLTELFKMKNAYEESLQOQLETLKRENKNLQOEISDLTEQTAEGGKRHLEBK 1532
QY 584 YKR--KQNELRLQGBELDKA 600
DB 1533 IKKQVEQEKSEIQAALAEA 1551

RESULT 10
Q9DGM5 PRELIMINARY; PRT; 1944 AA.
ID Q9DGM5;
AC Q9DGM5;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fast myosin heavy chain isoform 2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Q., Bandman E.;
RT "Seven skeletal myosin heavy chain genes (MYHC) are organized as a
RL multigene complex in the chicken genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF272033; AAF99314.1; -.
DR HSSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR SQ SEQUENCE 1944 AA; 223211 MW; 154F438220072D68 CRC64;
```

```
Query Match 13.5%; Score 418; DB 13; Length 1944;
Best Local Similarity 22.8%; Pred. No. 1.2e-07;
Matches 172; Conservative 139; Mismatches 244; Indels 200; Gaps 26;
```

```
QY 9 MKLRNKRFTKMSVMVKOE--GMELKQATQKDLTQESKGIQVLEGKLVSEIKSDEK 66
DB 838 MKLSFKIKPLKLSAESEKEMANMKEEFKTKELAKSAKRKELEKMMVTLKKNLQL 897
QY 67 ETEKLELYIQEISCASDQVEKCV-----DIALEEDLKEKREILSLKOSLEENIT- 118
DB 898 QVQAEADALAEERCQDLIKTKIQLEAKIKEVTERAEDEEINAEIATAKRRKLEDCSE 957
QY 119 FSKQIEDTVKQCLLETDRNLVSKDRERAETLSAEMQILTERALEROEYKQKQELQ 178
DB 958 LKKDIDDLTLAKYKEKHATENK----VKNLTEEMAGLDETIAKTKKALQEAHQ 1013
QY 179 SOSLQOQKE-----LSARLQOQLCSFOEEMTSEKNV----- 210
DB 1014 TLDDLQAEEDKYNVTUTKAKTKLEQVDDLEGSLEQEKLRMDLERAKRKLGLDLKMSQDT 1073
```

```
Qy 211 -----FKEELKIALAEALDAVQOEQESERLVKOLEEERKSTABQTRLDNLLREK 260
Db 1074 IMDLENDKQOLDKLKKDFEISQIQSIEDQALGMQLOKKIK--ELQARIEEL--EE 1128
Qy 261 EVELEKHIAHAQAAILIAQEKYN--DTAQSRLDVTVAQLESVQEKYNDTAQSRLDVTQ 319
Db 1129 EIEAERTSRKA-----EKHRADLSRELEFEISERLE---EAGGATATQI--DMNKKREA 1177
Qy 320 EOEKYNDTAQSRLDVTVAQLESVQEKYNDTAQSLR---DVTVAQLESVQEKYNDTAQSLRD 375
Db 1178 EFQKMR---RDLEATLQHEA-----TAAALRKKHADSTAEI-----GEQIDNLR 1220
Qy 376 VSAQLESYKSTLKEIEDLKENLTLOEKVAMAEKSVEDVQOQILTAESTNQEVARMVQD 435
Db 1221 VKOLEKESSELKMEIDDLASNMESVSKAKANLEKMCRTLEDQLSKIKSKEEHRQMIND 1280
Qy 436 LQ-NRSTLKEE-----EIKETSSFLKTTDLKNLRQ----- 468
Db 1281 LSTQARLQTESGYSRQVEEKDALISQLSRGKQAFQOIIEELKRLHLEETKAKNALAHA 1340
Qy 469 -----DED-FRKOLEEGKRTAEKNVMTLWINKWRLLY-----EELYEKTTP 513
Db 1341 LQSAHDCDLLREQYEEBQAKGLQALSKANSEVAQWRTKYETDAIQRTTEELERAKKK 1400
Qy 514 FQOQLD-----AFAEKQALLNE-----HGATQELNKRIRDSYAQL 549
Db 1401 LAQRLQDAEEHVEAVNAKASLEKTKQRLQNEVEDLMIDVERANACARLDKQKNFDKI 1460
Qy 550 LGHONLK-----OKKHVVYKLDENSQLSKSEVSKL 579
Db 1461 LAEWKQKYETQAELEASQESRSLSLSTELFKMKNAYEESLDHLETLKRENKNLQOEISDL 1520
Qy 580 RSOL-----VKR--KQNELRLQELDKA 600
Db 1521 TEQIAGGKAHLEKVKKQIEQEKSEIQAALAEA 1555

RESULT 11
Q9BB40
ID Q9BE40 PRELIMINARY; PRT; 1938 AA.
AC Q9BE40;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MYOSIN heavy chain 2x.
GN MYHC-2X.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HOLSTEIN; TISSUE-SKELETAL MUSCLE;
RA Chikuni K., Muroya S., Nakajima I.;
RT "Sequencing of the bovine myosin heavy chain isoforms.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB059399; BAB40921.1;
DR HSSP; P13538; 2MIS.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 1938 AA; 222988 MW; 95D05BE1B8978D40 CRC64;
```

```
Query Match 13.4%; Score 417; DB 6; Length 1938;
Best Local Similarity 23.0%; Pred. No. 1.3e-07;
Matches 170; Conservative 147; Mismatches 254; Indels 168; Gaps 24;

Qy 9 MKLNRKRETKMRSMVMKQOE--GMELKQATQKDTESKGKIVQLEKGLVSTIEKEKIDKCK 66
Db 833 KMLYFKIKPLLKSAETEEMANMKEFEKTEELAKSEAKRKELEKMTVTLTQEKNDLQL 892
Qy 67 ETEKLEVIQIISCASDOVEKCV-----DIAQLEEDLKEKOREILSLKOSLEENT- 118
Db 893 QVQSEADALADAEERCDQILIKTQLEAKIKEVTERAEDEEIEINAEIAKAKRKEDECE 952
Qy 119 FSKQIEDLTVCQQLLETERDNLVSKDRERAETLSAEMOILTERLALERQVEYKLOQKELQ 178
Db 953 LKKDIDDLLETLAKVEKEKHATENK---VKNLTTEAGLDETIAKLTKEKKALQEAHQ 1008
Qy 179 SQSLQOQE-----LSARLQOOLCSFQEMTSEKNV-----FKELKAL-A 220
Db 1009 TLDDLQAEEDKVNTLTAKTKLEQVDDLEGSLEQEKRLMDLERAKRKLKGLDLAOGES 1068
Qy 221 ELDAVQOEKESERLVKQ-----LE---EERKSTAEQTLRLDNLRLREKEVELEKHIAH 271
Db 1069 TMDIENDKQQDEKLKKKEFEMSNLQSKIEDEQALAMQKKIKELQARIIELEEEIEAE 1128
Qy 272 AQAIIAQEKYNDTAQSRLDVTVAQLESVQEKYNDTAQSRLDVTVAQLESVQEKYNDTAQSL 331
Db 1129 RASRAKAEKQSDLSRELEFEISERLE-----EAGGATSAQIEMNKKREAFQKMR 1178
Qy 332 RDVTAQLESVQEKYNDTAQSRLDVTVAQLESVQEKYNDTAQSRLDVSQALESYKSTLKEI 391
Db 1179 RD----LEEATLQHEATAAALRKKHA--DSVAE-LGEQIDNLRVKQKLEKSEKMKMEI 1231
Qy 392 EDLKENLTLOEKVAMAEKSVEDVQOQILTAESTNQEVARMVQDQONSTLKEEIKEIT 451
Db 1232 DDLASNMETVSKAGNLEKMCRALEDQLSELKTEDEQOQLNDLTQARLQTESGEFS 1291
Qy 452 -----SSFLEKTTDLKNLRQ-----DED-FRKOLE 477
Db 1292 RQLEKDALVSQLSRGKQAFQOIEELARQLEEEIKAKSALAHALQSAHDCDLLREQYE 1351
Qy 478 EKGKRTAEKENVMTLWINKWRLLY-----EELYEKTTPFQOQLD----- 519
Db 1352 EEQEGKAEQLQRAKSEKANSKSEVAQWRTKYETDAIQRTTEELERAKKLAQRLQDAEEHVEAVN 1411
Qy 520 ---AFAEKQALLNE-----HGATQEQIN-----KIRDSYAQL 550
Db 1412 AKCASLEKTKORLQNEVEDLMIDVERTNAACAALDKKORNFDKILSEWKQKYETHAELE 1471
Qy 551 GHQNLKQ-----KIKHWVK-----LKDENSQLSKSEKLSQL----- 583
Db 1472 ASQKESRSLSLSTELFKKNAYEESLDQLETLKRENKNLQOEISDLTEQIAGGKRIHELEK 1531
Qy 584 VKR--KQNELRLQELDKA 600
Db 1532 VKKQVEQEKSEIQAALAEA 1550

RESULT 12
Q14221
ID Q14221 PRELIMINARY; PRT; 1410 AA.
AC Q14221;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Endosome-associated protein.
GN EEA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RX MEDLINE=95286647; PubMed=7768953;
RA Mu F.T., Callaghan J.M., Steele-Mortimer O., Stenmark H., Parton R.G.,
RA Campbell P.L., McCluskey J., Yeo J.P., Tock E.P., Toh B.H.;
RT "EEA1, an early endosome-associated protein. EEA1 is a conserved
RT alpha-helical peripheral membrane protein flanked by cysteine
RT 'fingers' and contains a calmodulin-binding IQ motif.";
RL J. Biol. Chem. 270:13503-13511(1995).
DR EMBL; I40157; AAA79121.1; -.
DR InterPro; IPR000822; Znf_C2H2.
DR InterPro; IPR000306; Znf_FYVE.
DR Pfam; PF01363; FYVE; 1.
DR SMART; SM00064; FYVE; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 1410 AA; 162496 MW; DF0F9464D70A8AED CRC64;

Query Match 13.4%; Score 415; DB 4; Length 1410;
Best Local Similarity 22.8%; Pred. NO. 1.2e-07;
Matches 168; Conservative 153; Mismatches 223; Indels 192; Gaps 28;

QY 1 MRALSLEIMKLNR-----KRETKMRSMVYKQEG---MELK-----LQAT--QK 38
DB 283 VRVYVOELQKLKSSVNETQKNQITLNLKKEDYTKLEEKHNESVSKKNIQATLHQK 342
QY 39 DLTESGKGIVQLEGKLVSTKE-----KIDEKCE-TEKLLEYIOEISCASDQVECKVD 91
DB 343 DL-----DCQQLQSRLSASETSLSRIHVSEKGEATQKLEELSEV---ETKYQHLKAE 394
QY 92 IAQLEDLKEKREILSKQLEENITFSKQIEDLVTKCOLLETERDNLVSKDRERAETLSAE 151
DB 395 FKLOLQOQREKEQGLQLOSEINQ-----LHSLKLETERQLGEAHR-----L 437
QY 152 SAEQILTRLALEREQYEKLOQKELQSLLQKELSAKRLQOQL-----CSFQ 201
DB 438 KQROLSSBKMDKEQVADLQKLRLSEELKEKVTNSTELQHLQDKTKQOQHQEQALQ 497
QY 202 EEMTSKKNVKEELKALAEADVQKQESERLVKQ-----LEERK----- 244
DB 498 QSTAKLRAQNDLQVLRQIGCKEQKQINLEALLQKSKENISLLEKEDLYAKIQAGE 557
QY 245 ----STAQQLTRDNLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQE 301
DB 558 GETAVLNQLOEKHNTLQDQVTLQTEKLNQSESHKQAOENLHDQVQEQK---AHLRAAQD 614
QY 302 KYNDTAQSLRDVTAQLESQEKYND-----TAQSLRDVTAQLESQ 342
DB 615 RVLSTSVNELNSQLNESKEKVSQLDIQIKAKTELLLSAEAKTAQR-ADLNHLDTAQ 673
QY 343 EKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSTL---KETEDL---- 394
DB 674 NALQDKHQLNKITTLQDQVTKLQDKQHCQSLESHLKEYKEYLSLEQKTEELGQIK 733
QY 395 KLENLTQLKQVAMAEKSVEDVQOQILTAESTNQEYARMVQDQLNRSITLKEEBI---KEIT 451
DB 734 KLEADSLVK-ASKEQALQDLOQ-----RQNTDLELRATLSKQLEMEKEIV 781
QY 452 SSF-----LEKITYDLKNLRQOQDEDFRQLEBKGRKTAENVMTELMTWEINKRWLLY 504
DB 782 STRLDLQKKSEALSIEKQKLTQKEE--KQILKQDFETLSQET-----KIQH 827
QY 505 BELYEKTFFQOQLDAFAEAKQALLNEHCATQEQQLNKTRDSY-----AQL 549
DB 828 EELNNRIQTTVELQVKRMEKALMTSTLVKDKLSKVSQSDLSKSKSEFEKENQKGAIAI 887
QY 550 LGHQNLKQIKRHVVKLKDENS-----QLKSVSKLRSQLVKRR----- 586
DB 888 LDLEXTCKELKHQLOVQMENTLKEQELKKLSLEKEASHQLKLELNSMQEQLIAQNTL 947
QY 587 KON---ELRLQELDK 599
DB 948 KONEKEEQLOQGNINE 963

```

```

RESULT 13
Q15075 PRELIMINARY; PRT: 1411 AA.
ID Q15075
AC Q15075;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Endosomal protein.
GN P162.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Seelig H.P.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; X78998; CAA55632.1; -.
DR InterPro; IPR000822; Znf_C2H2.
DR InterPro; IPR000306; Znf_FYVE.
DR Pfam; PF01363; FYVE; 1.
DR SMART; SM00064; FYVE; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 1411 AA; 162465 MW; C3B17777FE34B6BD CRC64;

Query Match 13.3%; Score 414; DB 4; Length 1411;
Best Local Similarity 23.0%; Pred. NO. 1.3e-07;
Matches 168; Conservative 150; Mismatches 220; Indels 192; Gaps 28;

QY 7 ELMKURN-----KRETKMRSMVYKQEG---MELK-----LQAT--QKDLTSK 44
DB 289 ELQKLKSSVNETQKNQITLNLKKEDYTKLEEKHNESVSKKNIQATLHQKDL---- 344
QY 45 GKIVOLEKGLVSTKE-----KIDEKCE-TEKLLEYIOEISCASDQVECKVDIALE 97
DB 345 -DCQQLQSRLSASETSLSRIHVSEKGEATQKLEELSEV---ETKYQHLKAEFKQLOQ 400
QY 98 DLKEQILTRDNLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQEKYND 157
DB 401 QREKEQGLQLOSEINQ-----LHSLKLETERQLGEAHR-----LKEQRL 443
QY 158 LTERLALEREQYEKLOQKELQSLLQKELSAKRLQOQL-----CSFQEMTSE 207
DB 444 SSEKLMDEQOQVADLQKLRLSEELKEKVTNSTELQHLQDKTKQOQHQEQALQOQSTAK 503
QY 208 KNVKEELKALAEADVQKQESERLVKQ-----LEERK-----STA 247
DB 504 LREAQNDLQVLRQIGDKDQKQINLEALLQKSKENISLLEKEDLYAKIQAGEGAVL 563
QY 248 EQLTRDNLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTA 307
DB 564 NQLOEKHNTLQDQVTKLQDKQHCQSLESHLKEYKEYLSLEQKTEELGQIKLEADS 620
QY 308 QSLRDVTAQLESQEKYND-----TAQSLRDVTAQLESQEKYNDT 348
DB 621 TSVNELNSQLNESKEKVSQLDIQIKAKTELLLSAEAKTAQR-ADLNHLDTAQNALQDK 679
QY 349 AQSRLDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSTL---KETEDL----KLENLT 400
DB 680 QOELNKITTLQDQVTKLQDKQHCQSLESHLKEYKEYLSLEQKTEELGQIKLEADS 739
QY 401 LQEKVAMAEKSVEDVQOQILTAESTNQEYARMVQDQLNRSITLKEEBI---KEITSSF--- 454
DB 740 LEVK-ASKEQALQDLOQ-----RQNTDLELRATLSKQLEMEKEIVSTRLD 787
QY 455 ----LEKITYDLKNLRQOQDEDFRQLEBKGRKTAENVMTELMTWEINKRWLLYELVEK 510
DB 788 LQKKSEALSIEKQKLTQKEE--KQILKQDFETLSQET-----KIQEELNLR 833

```



SQ SEQUENCE 1939 AA; 223172 MW; B702ADB599602ECB CRC64;

Query Match 13.2%; Score 411; DB 6; Length 1939;  
 Best Local Similarity 22.7%; Pred. No. 2.2e-07;  
 Matches 168; Conservative 151; Mismatches 252; Indels 168; Gaps 25;

QY 9 MKLNRKRETKMRSMYKQE--GMELKQATQKDLTESKIVOLEGKLVISIEKIDKRC 66  
 ||| | : : : | : : | : : | : : | : : | : : | : : | : : | : :  
 Db 834 MKLYFKIKPILKSAETEKEMANNKEFEKTESLAKAEAKRKEEKVMVALMQEKNLQL 893  
 :  
 QY 67 ETEKLLLEYIQEISASDQVEKCKV-----DIAQLEEDLKEKDRILSILKQSLEENIT- 118  
 :  
 Db 894 QVOAEADSLADEERCQDLIKTIQIEAKIKEVTERAEDEEEINEAETAKKRKLEDECS 953  
 :  
 QY 119 FSKQIEDLVVKCOLLETERDNLYSKDRERAETLSAEMQILTERLALEROEYKLOQKELQ 178  
 | : | :  
 Db 954 LKKDIDDLLETLAKVEKEHATENK---VKNLITEEMAGUDETIAKLITREKALQEAHQ 1009  
 :  
 QY 179 SQSLLOQKE-----LSARLOQQLCSPQEWMTSEKNV-----FKBELKLAL-A 220  
 : || :  
 Db 1010 TLDDLQAEDKVNLTAKTKLEQVDDLEGSLEQEKLRMDLERAKRKLGDGDKLAQES 1069  
 :  
 QY 221 ELDAVQOKEEQSERLYKO-----LE---EERKSTAEQLTRLDNLLREKEVELEKHIAH 271  
 || :  
 Db 1070 TMDIENDKQOLDEKLLKKEFEMSNIQSKIEDEQALAMOLKKIKELQARIELEEIEAE 1129  
 :  
 QY 272 AQAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSL 331  
 :  
 Db 1130 RASRAKAEKQSDLSRELEISERLE-----EAGGATSAQIEMNKKREAEFOKMR 1179  
 :  
 QY 332 RDVTAQLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSL 391  
 || :  
 Db 1180 RD-----LEEATLQHEATAATLRKKHA--DSVAE-LGEQIDNLQRVKQKLEKEKSEMKMEI 1232  
 :  
 QY 392 EDLKLLENLTQEKVMAEKSVEDVQOQIILTAESTNQEYARMVQDL-QNRSTLKEE----- 445  
 || :  
 Db 1233 DLLASNMTVSKAGNLEKMCRTLEDQSELKTEEQOORLINDLTAQARLQTESGEYS 1292  
 :  
 QY 446 -----EIKETTSFLEKITDLKNOLRQO-----DED-FRKOLE 477  
 :  
 Db 1293 ROLDEKDTLVQSLSRCKQAFQOIEBELKQLEBEIKAKSAHAHVQSSRHDCDLLREQYE 1352  
 :  
 QY 478 EKGKRTAEKENVWELTMTINKWRLLY-----EELYEKTQPFQOQLD----- 519  
 | : | :  
 Db 1353 EEQEAQAELORAMSKANSEVAQWRTKYETDAIQRTTELEEAQKLAQRLQDAEEHVEAVN 1412  
 :  
 QY 520 ----APEAEKQALLNE-----HGATQEQBLN-----KIRDSYAQLL 550  
 : | | | | | :  
 Db 1413 AKCASLEKTKQRLQNEVEDLMIDVERSNAACAALDKORNFDKILAENWKQKYETHAELE 1472  
 :  
 QY 551 GHQNLKQ-----KIKHVVK-----LKDENSOLKSEVSKLRSQ----- 583  
 :  
 Db 1473 ASQKESRSLSSTELFKVKNAYEESLQOULETLKRENKNLQOEISDLTEQTAEGGKRIHELEK 1532  
 :  
 QY 584 VKR--KQNELRLOGELDKA 600  
 : | :  
 Db 1533 IKKQVEQEKSEIQAALAEA 1551

Search completed: December 20, 2002, 15:14:05  
 Job time : 46.6431 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:09:23 ; Search time 17.2176 Seconds  
(without alignments)  
3523.197 Million cell updates/sec

Title: US-09-685-010-48

Perfect score: 3104

Sequence: 1 MRALSLEMLKLRNKRRETKMR.....FCHASKENFTPLKEGPNPCC 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3098	99.8	631	2 JC4298	hyaluronan recepto
2	1953	62.9	725	1 JC5016	hyaluronan recepto
3	431.5	13.9	1790	2 S67593	transport protein
4	421	13.6	1938	2 A59293	skeletal myosin he
5	415	13.4	1410	1 A57013	early endosome ant
6	415	13.4	1940	1 A24922	myosin heavy chain
7	411	13.2	1938	1 JX0178	myosin heavy chain
8	408.5	13.2	1690	2 T13030	microtubule bindin
9	408	13.1	1130	2 T34081	hypothetical prote
10	404	13.0	1940	2 A29320	myosin heavy chain
11	397	12.8	1940	1 S04090	myosin heavy chain
12	393	12.7	1939	2 T48175	myosin heavy chain
13	391	12.6	2954	2 T14156	kinesin-related pr
14	390	12.6	1738	2 T14867	interaptin - slime
15	389.5	12.5	1935	1 A37102	myosin beta heavy
16	387	12.5	1938	1 S06005	myosin alpha heavy
17	387	12.5	1938	2 T49464	alpha cardiac myos
18	386	12.4	1392	2 A43336	microtubule-vesicl
19	384.5	12.4	2139	2 T18296	myosin heavy chain
20	384	12.4	1979	2 C71622	hypothetical prote
21	383	12.3	1427	2 S22695	restin - human
22	380.5	12.3	1931	2 A59234	slow myosin heavy
23	379.5	12.2	1937	2 T38055	myosin heavy chain
24	379.5	12.2	1938	1 A40997	myosin heavy chain
25	378.5	12.2	1934	2 T48153	myosin heavy chain
26	378	12.2	1837	2 T41023	probable nuclear p
27	377	12.1	1935	1 S06006	myosin beta heavy
28	376.5	12.1	978	2 A70387	conserved hypothet
29	376.5	12.1	1999	1 S21801	myosin heavy chain

30 375.5 12.1 1961 1 A61231 myosin heavy chain  
31 375.5 12.1 1976 2 A59252 myosin heavy chain  
32 375 12.1 1939 1 A46762 myosin alpha heavy  
33 374.5 12.1 2007 1 B43402 myosin heavy chain  
34 373 12.0 1909 2 A45592 liver stage antige  
35 373 12.0 1979 1 S03166 myosin heavy chain  
36 371.5 12.0 1935 2 A59286 364K Golgi complex  
37 371.5 12.0 3187 2 JC5837 KLP2 protein - Afr  
38 371 12.0 1388 2 T30335 giantin - human  
39 370 11.9 3225 2 T52300 giantin - human  
40 370 11.9 3259 1 A56539 giantin - human  
41 369 11.9 911 2 S51441 hypothetical prote  
42 367.5 11.8 1475 2 T33318 hypothetical prote  
43 366.5 11.8 1313 2 F96673 hypothetical prote  
44 366 11.8 1940 2 A59287 myosin heavy chain  
45 366 11.8 1972 1 A41604 myosin heavy chain

#### ALIGNMENTS

##### RESULT 1

JC4298

hyaluronan receptor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 16-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 01-Dec-2000

C:Accession: JC4298; A42925; A41923; S21586

R:Entwistle, J.; Zhang, S.; Yang, B.; Wong, C.; Li, Q.; Hall, C.L.; A, J.; Mowat, M.;

- Gene 163, 233-238, 1995

A:Title: Characterization of the murine gene encoding the hyaluronan receptor RHAMM.

A:Reference number: JC4298; MUID:96011639; PMID:7590272

A:Accession: JC4298

A:Molecule type: mRNA

A:Residues: 1-631 <ENT>

A:Cross-references: EMBL:X64550

A:Experimental source: 3T3 fibroblast

R:Hardwick, C.

- J. Cell Biol. 118, 753, 1992

A:Reference number: A42925; MUID:92348516; PMID:1639856

A:Contents: erratum

A:Accession: A42925

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 155-376, 'S', 378-504, 'E', 506-631 <HAR>

A:Cross-references: GB:X64550

A:Note: authors translated the codon AGT for residue 377 as Thr and CTA for residue 5

R:Hardwick, C.; Hoare, K.; Owens, R.; Hohn, H.P.; Hook, M.; Moore, D.; Cripps, V.; Au

- J. Cell Biol. 117, 1343-1350, 1992

A:Title: Molecular cloning of a novel hyaluronan receptor that mediates tumor cell mo

A:Reference number: A41923; MUID:92299690; PMID:1376732

A:Accession: A41923

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 155-376, 'S', 378-504, 'E', 506-507, 'L', 508-630 <HA2>

A:Cross-references: GB:X64550

A:Note: this sequence has been corrected in reference A42925

C:Comment: This protein regulates cell motility and transformation, and focal adhesio

C:Genetics:

A:Gene: rhamm

A:Introns: 20/3; 54/2; 79/2; 137/3; 187/3; 259/2; 382/2; 431/2; 482/2; 515/3; 574/3;

C:Superfamily: hyaluronan receptor

C:Keywords: glycoprotein; receptor

F:260-382/Region: 21 residue repeats

F:516-574/Region: hyaluronan binding #status predicted

F:575-625/Region: hyaluronan binding #status predicted

F:116,283,304,325,346,367,398,438,619/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 99.8%; Score 3098; DB 2; Length 631;

Best Local Similarity 99.7%; Pred. No. 3.3e-100;

Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRALSLEMLKLRNKRRETKMRSMVKQEGMELKQATKDLTSGKIVQLEKGLVSIKKE 60

|||||

Db 1 MRALSLEMLKLNKRETKRMVMVQEGMELKLOATQKDLTSGKIVQLEGKLSIEKE 60  
QY 61 KIDEKCEKLELEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENIT-F 120  
Db 61 KIDEKCEKLELEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENIT-F 120  
QY 121 KQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKIQKQELQSQ 180  
Db 121 KQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKIQKQELQSQ 180  
QY 181 SLLOQEKELSAEQOOLCSFOEEMTSKKNVFEELKLAELAELDAVQOKEQSERLVKQLE 240  
Db 181 SLLOQEKELSAEQOOLCSFOEEMTSKKNVFEELKLAELAELDAVQOKEQSERLVKQLE 240  
QY 241 EERKSTAEOLTRDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300  
Db 241 EERKSTAEOLTRDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300  
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360  
Db 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360  
QY 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420  
Db 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 420  
QY 421 TAESTNOEYARMVQDLQNRSTLKEEIEKITSFLEKITDKNLQROODEDFRKOLEEGK 480  
Db 421 TAESTNOEYARMVQDLQNRSTLKEEIEKITSFLEKITDKNLQROODEDFRKOLEEGK 480  
QY 481 KRTAEKENVMTLMEINKWRLLYELYEKTKPFQOQDLDAFAEAKOALLNEHGATQEQLN 540  
Db 481 KRTAEKENVMTLMEINKWRLLYELYEKTKPFQOQDLDAFAEAKOALLNEHGATQEQLN 540  
QY 541 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLSRQKONELRLQLEGDKA 600  
Db 541 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLSRQKONELRLQLEGDKA 600  
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631  
Db 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631

RESULT 2  
JC5016  
hyaluronan receptor - human  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 29-Sep-1999  
C:Accession: JC5016  
R:Wang, C.; Entwistle, J.; Hou, G.; Li, Q.; Turley, E.A.  
Gene 174, 299-306, 1996  
A:Title: The characterization of a human RHAMM cDNA: Conservation of the hyaluronan-binding site in the human RHAMM protein  
A:Reference number: JC5016; MUID:97045829; PMID:8890751  
A:Contents: breast  
A:Accession: JC5016  
A:Molecule type: mRNA  
A:Residues: 1-725 <W>  
A:Cross-references: GB:U29343  
A:Note: it is uncertain whether Met-1 or Met-196 is the initiator  
C:Comment: This receptor regulates focal adhesion turnover, and regulates human breast cancer cell growth  
C:Genetics:  
A:Gene: GDB:HMMR; RHAMM  
A:Cross-references: GDB:683209; OMIM:600936  
A:Map position: 5q33.2-5qter  
C:Superfamily: hyaluronan receptor

Query Match 62.9%; Score 1953; DB 1; Length 725;  
Best Local Similarity 64.8%; Pred. No. 9.7e-61;  
Matches 411; Conservative 49; Mismatches 86; Indels 88; Gaps 3;

QY 1 MRALSLEMLKLNKRETKRMVMVQEGMELKLOATQKDLTSGKIVQLEGKLSIEKE 60  
Db 164 LRILSLEMLKLNKRETKRMVMVQEGMELKLOATQKDLTSGKIVQLEGKLSIEKE 223

QY 61 KIDEKCEKLELEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENIT-F 119  
Db 224 KIDEKCEKLELEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENIT-F 283  
QY 120 KQIEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKIQKQELQSQ 179  
Db 284 SKQVEDLVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKIQKQELQSQ 343  
QY 180 QSLLOQEKELSAEQOOLCSFOEEMTSKKNVFEELKLAELAELDAVQOKEQSERLVKQLE 239  
Db 344 DLSLLOQEKELSAEQOOLCSFOEEMTSKKNVFEELKLAELAELDAVQOKEQSERLVKQLE 403  
QY 240 EERKSTAEOLTRDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 299  
Db 404 EERKSTAEOLTRDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 457  
QY 300 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 359  
Db 458 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 457  
QY 360 EKVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 419  
Db 458 EKVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 499  
QY 420 TAESTNOEYARMVQDLQNRSTLKEEIEKITSFLEKITDKNLQROODEDFRKOLEEGK 479  
Db 500 LATESNOEYARMVQDLQNRSTLKEEIEKITSFLEKITDKNLQROODEDFRKOLEEGK 559  
QY 480 KRTAEKENVMTLMEINKWRLLYELYEKTKPFQOQDLDAFAEAKOALLNEHGATQEQOL 539  
Db 560 EERKSTAEOLTRDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESV 619  
QY 540 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLSRQKONELRLQLEGDK 599  
Db 620 KIRDSYAQLLGHONLKQKIKHVVKLKDENSOLKSEVSKLSRQKONELRLQLEGDK 679  
QY 600 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 630  
Db 680 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 713

RESULT 3  
S67593  
transport protein USO1 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein D2552; protein YDL058w  
C:Species: Saccharomyces cerevisiae  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: S67593; A38455; S30782  
R:Blöcker, H.; Brandt, P.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67597  
A:Accession: S67593  
A:Molecule type: DNA  
A:Residues: 1-1790 <BLO>  
A:Cross-references: EMBL:Z74106; NID:g1431058; PID:e253003; PID:g1431059; MIPS:YDL058  
A:Experimental source: strain S288C  
R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.  
J. Cell Biol. 113, 245-260, 1991  
A:Title: A cytoskeleton-related gene, USO1, is required for intracellular protein trafficking  
A:Reference number: A38455; MUID:91185402; PMID:2010462  
A:Accession: A38455  
A:Molecule type: DNA  
A:Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>  
A:Cross-references: GB:X54378; NID:g4777; PIDN:CAA38253.1; PID:g4778  
A:Note: the authors translated the codon ACT for residue 768 as Ile  
R:Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.  
submitted to the EMBL Data Library, February 1993  
A:Description: An integrin analogue in Saccharomyces cerevisiae.  
A:Reference number: S30782  
A:Accession: S30782  
A:Molecule type: DNA  
A:Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580



A;Cross-references: EMBL:L03188

C;Genetics:

A;Gene: SGD:USO1; INT1

A;Cross-references: SGD:S0002216; MIPS:YDL058W

A;Map position: 4L

C;Keywords: coiled coil; transmembrane protein

F;326-342/Domain: transmembrane #status predicted &lt;TMI&gt;

F;394-410/Domain: transmembrane #status predicted &lt;TM2&gt;

F;617-633/Domain: transmembrane #status predicted &lt;TM3&gt;

Query Match 13.9%; Score 431.5; DB 2; Length 1790;

Best Local Similarity 24.2%; Pred. No. 5.5e-08;

Matches 179; Conservative 149; Mismatches 236; Indels 177; Gaps 33;

Qy 10 KLRNKRKTRKMRMMVKQEGMELKQAYQKOLTESKGIQVLEGKLVSIKKEKIDKCE 69

Db 798 ELKNVRDS-LDEMQLRDVLTEDKQNTALLEYKSTIHQEDSIKTLEK----- 846

Qy 70 KLEYIQEISCASDOVECKVDIAQLEEDL-----KEKD-----REILSLKQ 111

Db 847 GLETILSQKKAEDGINKMGKDLFALSRMQAVEENCKLQKEDKSNVNHQKETSLSKE 906

Qy 112 SLEENIT---FSKOIEDLVTKCOLLETER-----DNLVSKDRERAET 150

Db 907 DIAAKITEIKAINLENLEEMKIQCNNLSKEKEHISKELVEYKSRFSQSHDNVLAKITELKLS 966

Qy 151 LS---AEMQILTERL--ALEROEYE-KLQKELQSQ-SLLQOEKELSLARLOOQLCSFOEE 203

Db 967 LANNYKDMAQENESLIKAVEKSNESSTOLSNLKNKIDSNQSEKE-----NQIE 1016

Qy 204 MTS-EKNVFEELKALAEELDAVOOK-----EEOSERLVQLEERKSTAPQL 250

Db 1017 RGSIEKNI--EQLKTTISDLEQTEETIISKSDSKDEYVESQISLLKEKLETATTANDNV 1074

Qy 251 TRLDNLLREKEVELEKHTAAHAQAILTAQEKYNTAQSRLDVTQALESVQE-----KY 303

Db 1075 NKISLKTRE-ELEAELAAYNKLNKNELETSETKALKEVNEEHLKKEKIQLEKEA 1133

Qy 304 NDTAQSRLDVTQALESVQEKYNDTAQSRLDVTQALESVQEKYNDTAQSRLD-VTA---QL 359

Db 1134 TETQOQLNSLRANLESLEKEHEDLAQULKVYEOIANKERYNEEISQNDNEITSTQOEN 1193

Qy 360 ESVOEKYND---TAQSRLDVSQLESYKSTLK-----EIEDLKLENLT----- 400

Db 1194 ESIRKKNDELEGEYKAMKSTSEQSNLKKSEIDALNLQIKELKKKNETNEASLESISKV 1253

Qy 401 -----LQEKYAMAESVEDVQOQILTAESTNQEYARMVQDLONRSTLKEEIKEIT 451

Db 1254 ESETVKIKELQDECNFKEVESELEDKLKASEDNKSKYL-----ELQKESEKIKBELDAKT 1309

Qy 452 SSF---LEKITDLKNLRQODEDFRQOLEEKGRKTAKENVMTELTMINKWRLLYEELY 508

Db 1310 TELAIQLEKITNL-----SKAKESESE---LSRLKKTSSERNNAEBQL 1351

Qy 509 EKTQ-PFQOQLDAFAEKQALLNEHGAT-----QEQLNKRDSYAQLGHQNLKQIKHV 562

Db 1352 EKLNEIQIKNOAQEKERK-LLNQSSTITOYESEKINTLDEILRLQENELKAK----- 1406

Qy 563 VKLDENSOLKSEYKSLRSQLVKQNELR-LOQEL-----DKALGI-----RH 605

Db 1407 -EIDNTRSELE-KVSLSNDELLEKQNTIKSLQLODEILSYKDKITRNDEKLLSIEDRNKR 1464

Qy 606 FDPKAFCHASKENFTPLKEG 626

Db 1465 LESLKEQLRAAQESKARVEEG 1485

RESULT 4

A59293

skeletal myosin heavy chain - domestic rabbit

C;Species: Oryctolagus cuniculus

C;Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 08-Sep-2000

C;Accession: A59293

R;Maeda, K.; Hostinova, E.; Roesc.Kleinkauf, A.; Schuster, H.; Gasperik, J.; Wittingh submitted to GenBank, July 1995

A;Description: Isolation, sequencing of myosin heavy chain cDNA from rabbit skeletal

A;Reference number: A59293

A;Accession: A59293

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1938 &lt;MAE&gt;

A;Cross-references: GB:U32574; NID:g940232; PIDN:AAA74199.1; PID:g940233

A;Experimental source: strain New Zealand White; cell type skeletal muscle fiber type

C;Genetics:

A;Gene: MHC

C;Superfamily: myosin heavy chain; myosin motor domain homology

F;89-769/Domain: myosin motor domain homology &lt;MWO&gt;

Query Match 13.6%; Score 421; DB 2; Length 1938;

Best Local Similarity 23.0%; Pred. No. 1.4e-07;

Matches 170; Conservative 150; Mismatches 251; Indels 168; Gaps 25;

Qy 9 MCLRNRKTRKMRMMVKQEGMELKQATQKDLTESKGIQVLEGKLVSIKKEKIDKCE 66

Db 833 MKLVFKIRPLLKSAETEKEMANMKKEFEKTLAKAKAEKELEKVMALMQEKNDLQL 892

Qy 67 ETEKLEYIQEISCASDOVECKV-----DIAQLEEDLKEKDEILSLKOSLEENIT- 118

Db 893 QVQAEADSLADAEERCDOILIKIQLEAKIKVEVTERAEDEEINAEITAKKRKLEDECE 952

Qy 119 FSKQIEDLVTKCOLLETERDNLVSKDRERAETLSAEMOILTERLALEHQEYKLOQKELO 178

Db 953 LKKDIDDLLELTIAKVEKEKHATEN-----VKNLTTEMAGLDETIAKLTKEKALQEAHQO 1008

Qy 179 QSLSLQBEKE-----LSARLOOQLCSFOEEMTSEKNV-----FKBELKAL-A 220

Db 1009 TLDDQLAEEDKVNTLTAKTKLEQQVDDLEGSLEQKIRMDLERAKRKLEGDLKLAQES 1068

Qy 221 ELDAVQOKEEESERLVKQ-----LE---EERKSTAEQILTRLDNLLREKEVELEKHIAAH 271

Db 1069 TMDIENDRQOQLDEKLKKEFEFENLSQKIEDQALAMOLQKKIKELQARIEELEEEIEAE 1128

Qy 272 AQAILTAQEKYNDTAQSRLDVTQALESVQEKYNDTAQSRLDVTQALESVQEKYNDTAQSL 331

Db 1129 RASRAKAKQSDJUSLEEEISERLE-----EAGGATSAQIEMNKKREAEFQKMR 1178

Qy 332 RDVTAQLESQEKYNDTAQSRLDVTQALESVQEKYNDTAQSRLDVSQLESYKSTLKEI 391

Db 1179 RD---LEEATLQHEATAATLRKHA--DSVAE-LGEQIDNLRQVKQLEKESEKELKWEI 1231

Qy 392 EDLKLENLTQEKYVAMAESVEDVQOQILTAESTNQEYARMVQDLO-NRSTLKEE----- 445

Db 1232 DLLASNMTVSKAKGNLEKMCRTLEDQVSELKTEEEHQRLINDLSAORARLQTESGEFS 1291

Qy 446 -----EIKETSSFELEKITDLKNLRQO-----DED-FRQOLE 477

Db 1292 ROLDEKDSLVSGKQGAFTQOIEELKRQLEEEITAKSALAHALQSAHRHCDLLREQYE 1351

Qy 478 EKGKRTAEKENVMTELTMEINKWRLLY-----EELYEKTQPFQOQLD----- 519

Db 1352 EEQAKAELQRAMKANSEVAQWRKYETDAIQRTTEEELEEAKKLAQRLQDAEEHVEAVN 1411

Qy 520 ----AFAEKQALUNE-----HGATQEQNL-----KTRDSYAQLL 550

Db 1412 AKCASLEKTKORLQNEVEDLMIDVERTNAACAALDKQKQNFQKILAEWKHKYEETHAELE 1471

Qy 551 GHQNLKQ-----KIKHVVK-----LKDENSOLKSEVSKLSQL----- 583

Db 1472 ASQKESRSLSTEVFKVKNAYEESLDQLETLKRENKNLQOETSIDLTEQIABEGKRIHELEK 1531

Qy 584 VKR--KQNELRLQGEELDKA 600

Db 1532 VKKQVEQEKSELQAALAEA 1550

RESULT 5

A57013  
nearly endosome antigen 1 - human  
N:Alternate names: endosome-associated protein  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A57013; S44243  
J. Biol. Chem. 270, 13503-13511, 1995  
R. Mu, F.T.; Callaghan, J.M.; Steele-Mortimer, O.; Stenmark, H.; Parton, R.G.; Campbell, J.  
A:Title: EEAL1, an early endosome-associated protein. EEAL1 is a conserved alpha-helical p  
A:Reference number: A57013; MUID:95286647; PMID:7768953  
A:Accession: A57013  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1410 <RES>  
A:Cross-references: GB:L40157; NID:g1016367; PIDN:AAA79121.1; PID:g1016368  
R:Seelig, H.P.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: S44243  
A:Accession: S44243  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-254, 'C', 256-257, 'LQ', 260-276, 'A', 278-283, 'A', 285-519, 'D', 521-574, 'EQ', 577-  
A:Cross-references: EMBL:X78998; NID:9475933; PIDN:CAA55632.1; PID:9475934  
C:Genetics:  
A:Gene: GDB:EEAL  
A:Cross-references: GDB:1369996  
C:Superfamily: human early endosome antigen 1  
C:Keywords: calmodulin binding; endocytosis; metal binding; peripheral membrane protein;  
  
Query Match 13.4%; Score 415; DB 1; Length 1410;  
Best Local Similarity 22.8%; Pred. No. 1.6e-07;  
Matches 168; Conservative 153; Mismatches 223; Indels 192; Gaps 28;  
  
Qy 1 MRALSLEMLKLN-----KRETKRSMVMVKQEG---WELK-----LOAT--OK 38  
Db 263 VRYVYQELQKSSVNETQKNTLTENLLKKEQDYTKLEKHNEESVSKNIIQATLHQK 342  
Qy 39 DITSGKXIVQLEGLKLVSIKE-----KIDKCB-TEKLLIYQIISCASDQVEKCKVD 91  
Db 343 DL-----DCQQLQSLASSETSLSRIHVSELSKGEATQKLEELSEV---ETKYQHLKAE 394  
Qy 92 IQLBEDLKDREILSLKSLSEENITFSKQIEDITVVCQQLLETDRDNLVSKDRERATL 151  
Db 395 FKLOQQQREKQHQGLQSEINQ-----LHKKLETERQLGEAHR-----L 437  
Qy 152 SAEMQILTERLALEREQYEKQKELQSQSLQOEKLSARLQOOL-----CSFQ 201  
Db 438 KEQRLSSEKLMDEQVADQLKLSRLEEQLKEKVTNSTELOHQLDKTKQHQQEQALQ 497  
Qy 202 EEMTSEKNVFEELKALAEILDVQOEQSERLVKQ-----LEERK----- 244  
Db 498 QSTTAKLREAQNDLQVLRQIGEKDQKIQNLLEALLQSKENISLLEKEREIDYAKIQAGE 557  
Qy 245 ---STAEQLTRLDNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQE 301  
Db 558 GETAVNLQEKNHPTLQDEVTLQNTENVNQNQSESHKQAQENLHDQVQEQK---AHLRAAQD 614  
Qy 302 KYNDTAQSLRDVTAQLESQEKYND-----TAQSLRDVTAQLESQ 342  
Db 615 RVLSLTSVNLNSQLNSKESKVSQSLDIQIKAKTELLLSABAATAQR-ADLNHLDTAQ 673  
Qy 343 EKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSSITL-----KEITDL----- 394  
Db 674 NALQDKHQELNKIITQLDQVNTAKLQDKQEHCSQLESHLKEYKLYLSLEQKTEELGOIK 733  
Qy 395 KLENTLQEKVAMAKSVEDVQQQIILTAESTNQEVARMVQDLQNRSTLKEBEI---KEIT 451  
Db 734 KLEADSELVK-ASKERQALQDLQQQ-----RLNTDLRLATLSKQLEMEKEIV 781  
Qy 452 SSF-----LEKITDLKNLRQOODEFRKOLEEGKGTAEKNVMTLTWTWINKRWLLY 504  
Db 782 SSTRLDLQKKSEALSISQKUTKQEE--KQILKQDFTLTQET-----KIQH 827

Qy 505 EELYEKTRPFOOQDLDAFEAEKQALLNEHGATQEQINKIRDSY-----AQL 549  
Db 828 EELNNRIQTIVTELQVKMEKEALMTLSTVTKDKLSKYSKSEFEKENQKGAAI 887  
Qy 550 LGHQNLKQKIKHVVLKDENS-----QLKSEVSKLRSOLVVR----- 586  
Db 888 LDLEKTCRELKHQLOQVOMENTLKEQKELKSLSEKEASHQKLENSMQEQLIQANTL 947  
Qy 587 KON---ELRLQGELEDK 599  
Db 948 KQNEKEEQLOQNGINE 963  
  
RESULT 6  
A24922  
myosin heavy chain, skeletal muscle, embryonic - rat  
N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 19-Apr-2002  
C:Accession: A24922; A22538; B24263  
R:Strehler, E.E.; Strehler-Page, M.A.; Perriard, J.C.; Perlasamy, M.; Nadal-Ginard, B  
J. Mol. Biol. 190, 291-317, 1986  
A:Title: Complete nucleotide and encoded amino acid sequence of a mammalian myosin he  
A:Reference number: A24922; MUID:87060988; PMID:3783701  
A:Accession: A24922  
A:Molecule type: DNA  
A:Residues: 1-1940 <STR>  
A:Cross-references: GB:X04267; GB:X05004; NID:g56658; PIDN:CAA27817.1; PID:g1619338  
R:Strehler, E.E.; Mahdavi, V.; Perlasamy, M.; Nadal-Ginard, B.  
J. Biol. Chem. 260, 468-471, 1985  
A:Title: Intron positions are conserved in the 5' end region of myosin heavy-chain ge  
A:Reference number: A22538; MUID:85080119; PMID:2981212  
A:Accession: A22538  
A:Molecule type: DNA  
A:Residues: 1-168 <STR>  
A:Cross-references: GB:L00370; GB:M0135; NID:g205580; PIDN:AAA41655.1; PID:g9554476  
R:Perlasamy, M.; Wydro, R.M.; Strehler-Page, M.A.; Strehler, E.E.; Nadal-Ginard, B.  
J. Biol. Chem. 260, 15856-15862, 1985  
A:Title: Characterization of cDNA and genomic sequences corresponding to an embryonic  
A:Reference number: A24263; MUID:86059474; PMID:2999140  
A:Accession: B24263  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1358-1490, 'G' <PER>  
A:Cross-references: GB:K03468; NID:g205573; PIDN:AAA41652.1; PID:g205574  
A:Experimental source: clone pMHC-72  
C:Genetics:  
A:Introns: 68/3; 116/3; 169/1  
A:Note: the list of intron positions may be incomplete  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle  
F:89-767/Domain: myosin motor domain homology <MMOT>  
F:179-186/Region: nucleotide-binding motif A (P-loop)  
F:549-586/Region: actin binding #status predicted  
F:656-678/Region: actin binding #status predicted  
F:840-1940/Domain: coiled coil #status predicted <COI>  
F:840-1280/Region: S2  
F:1281-1940/Region: light meromyosin  
F:130/Modified site: N6,N6-trimethyllysine (Lys) #status predicted  
F:185/Binding site: ATP (Lys) #status predicted  
F:696,706/Active site: Cys #status predicted  
  
Query Match 13.4%; Score 415; DB 1; Length 1940;  
Best Local Similarity 22.0%; Pred. No. 2.2e-07;  
Matches 165; Conservative 144; Mismatches 244; Indels 198; Gaps 22;  
  
Qy 4 LSLEMLKLNRRKRETKRSMVMVKQEGMELKLAQTKDLPESKGIIVQLEKGLVSTIEKEID 63  
Db 833 LFFKIKPLKLSAETKEKEMATMKEE-----FQTKDELAKSEAKRKEELKLVTLVQEKND 887  
Qy 64 ----EKCEKLELYIQEISCASDOVEKCV-----DIAQLEDLKEKDEILSLKQS 112  
Db 888 LQVQQAENSENLID--AEERC--DQLIKAKFQLEAKIKVEVTERAFDEEINAEITAKRRK 943







Db 895 QVOAEDAGLADAEBCDQIKTKIQLEAKIKELTERAEDEEENNAELTAKKKRLEDECSE 954  
QY 119 FSKQIEDLVTKCOLLETERDNLVSKDRERAETLSAEMOILTERLALERQYERKQLQKELQ 178  
Db 955 LKKDIDDLLETLAKVEKHEATENK----VKNLTTEMAALDETIAKTREKRALQEAHQO 1010  
QY 179 S-----QSLLOQKEL-----SARLOOOLCSFQB- 202  
Db 1011 TLDDLAQEEEDKVNLTAKTKLEQQVDDLEGSLEQEKKLMDLERAKKRLEGDLKMTQES 1070  
QY 203 --EMTSEKNVFEELKLAELDAVQOKEEQSERLVKQLEEBKSTAEQTLRLDNLRLRK 260  
Db 1071 TMDLEKDQOLDEKLKKDFEISQIQSKIEDEQALGMOLKKIK----ELQARIEEL--EE 1125  
QY 261 EVELEKHIAHAQAAILIAOEKYN-DTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLES 319  
Db 1126 ETEAETSRAK-----EKHRADLSRELEISERLE---FAGGATAAQI-DMKKKRA 1174  
QY 320 EQEKYNDTAQSLRDVTAQLESE----QEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRD 375  
Db 1175 EFQKMR---RDLEAATLQHEATAAALRKHK---ADSTADVGEQIDNLQR----- 1217  
QY 376 VSAQLESYSSTLKETEDLKENLTLQEKVMAEK---SVEDVQOQILTAESTNQ----- 427  
Db 1218 VKQLEKESKELMETDDLASNMESVSKANLEKMCRLSDQLSEIKTKEEQOQTIND 1277  
QY 428 -----EYARMVQDLQNRSTLKEEIKETT---SSFLEKITDLKKNLQOQ--- 468  
Db 1278 ISAKARLOTESGEYSRQVEE-----KDALISLSRGKQAFQTQOIELKRLHLEEIK 1330  
QY 469 -----DED-FRKQLEEKGRTPAKENVMTELTMEINKWRLLY-----EE 506  
Db 1331 KKCPAHALQASRHDCDLLREQVEEQEAKGELQALSKANSEVAQWRTKYETDAIQRTTE 1390  
QY 507 LYEKTRPFQOQLD-----APEAEKQALLNE-----HGATQOLNKI 542  
Db 1391 LEAKKKLQARLQDAEBHVEAVNSKASLEKTQRQLQNEVEDLMDIVERSNAACAALDKK 1450  
QY 543 RQSYAQLLGHQNLK-----VVR--KQNELRLQGLDKA 600  
Db 1451 QKNFDKILSEWKQKYEETQAELEASQESRSLSLETLFKMKNAYESLDHLETLKRENKL 1510  
QY 573 KEVSKLRSQ-----VVR--KQNELRLQGLDKA 600  
Db 1511 QOEISDLTEQIAEGGKAIHELEKVKKKIQBEKSELTQALEEA 1552  
RESULT 11  
S04090  
myosin heavy chain 3, skeletal muscle, embryonic - human  
N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 19-Apr-2002  
C:Accession: S04090; S06146; S05442; S12460; S09333; A35082  
R:Eller, M.; Stedman, H.H.; Sylvester, J.E.; Fertels, S.H.; Rubinstein, N.A.; Kelly, A.M.  
Nucleic Acids Res. 17, 3591-3592, 1989  
A:Title: Nucleotide sequence of full length human embryonic myosin heavy chain cDNA.  
A:Reference number: S04090; MUID:89263803; PMID:2726495  
A:Accession: S04090  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-1940 <ELL>  
A:Cross-references: EMBL:X13988; NID:g34843; PIDN:CAA32167.1; PID:g34844  
R:Eller, M.; Stedman, H.H.; Sylvester, J.E.; Fertels, S.H.; Wu, Q.L.; Raychowdhury, M.K.  
FEBS Lett. 256, 21-28, 1989  
A:Title: Human embryonic myosin heavy chain cDNA. Interspecies sequence conservation of  
A:Reference number: S06146; MUID:90033298; PMID:2806546  
A:Accession: S06146  
A:Molecule type: mRNA  
A:Residues: 774-1662, 'OT', 1665-1940 <EL2>  
A:Cross-references: EMBL:X13100; NID:g31143; PIDN:CAA31492.1; PID:g31144  
R:Karsch-Mirzachi, I.; Travis, M.; Blau, H.; Leinwand, L.A.  
Nucleic Acids Res. 17, 6167-6179, 1989

A:Title: Expression and DNA sequence analysis of a human embryonic skeletal muscle my  
A:Reference number: S05442; MUID:89366648; PMID:2771643  
A:Accession: S05442  
A:Molecule type: DNA  
A:Residues: 856-1390, 'KK', 1393-1940 <KAR>  
A:Cross-references: EMBL:X15696; NID:g36504; PIDN:CAA33731.1; PID:g1335313  
R:Stedman, H.H.; Eller, M.; Jullian, E.H.; Fertels, S.H.; Sarkar, S.; Sylvester, J.E.  
J. Biol. Chem. 265, 3568-3576, 1990  
A:Title: The human embryonic myosin heavy chain. Complete primary structure reveals e  
A:Reference number: A35082; MUID:90154023; PMID:2303463  
A:Contents: annotation; chromosomal assignment  
R:Bober, E.  
submitted to the EMBL Data Library, January 1989  
A:Reference number: S12458  
A:Accession: S12458  
A:Molecule type: mRNA  
A:Residues: 856-1330, 'G', 1332-1390, 'KK', 1393-1607, 'RA', 1610-1940 <BOB>  
A:Cross-references: EMBL:X51593; NID:g29463; PIDN:CAA35942.1; PID:g29464  
A:Experimental source: clone gTMC-E  
R:Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.  
Eur. J. Biochem. 189, 55-65, 1990  
A:Title: Identification of three developmentally controlled isoforms of human myosin  
A:Reference number: S09331; MUID:90235862; PMID:1691980  
A:Accession: S09333  
A:Molecule type: mRNA  
A:Residues: 856-901, 'X', 903-971, 'X', 973-1041, 'X', 1043-1111, 'X', 1113-1181, 'X', 1183-125  
1, 'X', 1673-1741, 'X', 1743-1811, 'X', 1813-1881, 'X', 1883-1940 <BOB>  
A:Cross-references: EMBL:X51593  
C:Genetics:  
A:Gene: GDB:MYH3  
A:Cross-references: GDB:119443; OMIM:160720  
A:Map position: 17p13.1-17p13.1  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle  
F:89-767/Domain: myosin motor domain homology <MMOT>  
F:179-186/Region: nucleotide-binding motif A (P-loop)  
F:549-586/Region: actin binding #status predicted  
F:656-678/Region: actin binding #status predicted  
F:840-1940/Domain: coiled coil #status predicted <COI>  
F:840-1280/Region: S2  
F:1281-1940/Region: light meromyosin  
F:130/Modified site: N6,N6, trimethyllysine (Lys) #status predicted  
F:185/Binding site: ATP (Lys) #status predicted  
F:696,706/Active site: Cys #status predicted  
Query Match 12.8%; Score 397; DB 1; Length 1940;  
Best Local Similarity 21.5%; Pred. No. 9.2e-07;  
Matches 165; Conservative 143; Mismatches 258; Indels 200; Gaps 22;  
QY 4 LSLELMKLRNKRTRKMSVMYKQGMELKQATQKDLTSKGIQVLEGKLVSIKREID 63  
Db 833 LFFKIKPLKLSAETKEKATMKEE----FQTKDELAKSEAKRKELEKLVTLVQEKND 887  
QY 64 ----EKCTEKLELYIQEISASDOVEKCV-----DIAQLEDLKEKREILSLKQS 112  
Db 888 LQLQVQAESENLDD--AEERC--DOLIKAFQLEAKIREVTERAEDEEINAEALTAKRK 943  
QY 113 LEENIT-FSKOIEDLVTKCOLLETERDNLVSKDRERAETLSAEMOILTERLALERQYERK 171  
Db 944 LEDECSLEKLDIDDLLETLAKVEKHEKHAENKVNKLTELSELSG----LDETAKLTREKA 999  
QY 172 LQOELQSQSLQOKE-----LSARLOOOLCSFQEMTSEKNV-----FKEE 214  
Db 1000 LQEAHQALDDLOAEEDKVNLSLTKSKLEQOVEDLESSLEQEKLRVLDERNKRKLEGD 1059  
QY 215 LKLALAE-LDAVQOKEEQSERLVK-----QLE----EERKSTAEQTLRLDNLRLKEVEL 264  
Db 1060 LKLAQESTLIDLENDKQQLDERLKKDFCYQLQSKVEDEQTLGLQFQKKELQARIEL 1119  
QY 265 EKHTAAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEQ-K 323  
Db 1120 EETEAERATRAKTEKQKRSQSDYARELESELSERLEE-----AGGVTTQTIELNKKRAE 1171











Db 1461 EESQSESSQKDEARSLSTELFKLNAYEESLEHLETFKRENKNLQOEISDLTEQLGSSG 1520  
QY 586 -----RKONE---LRLQGLDKA 600  
Db 1521 KTIHELEKVRKQLEAEKMELOSALEA 1547

Search completed: December 20, 2002, 15:12:55  
Job time : 34.2176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:06:33 ; Search time 32.5737 Seconds  
(without alignments)  
2581.255 Million cell updates/sec

Title: US-09-685-010-48  
Perfect score: 3104  
Sequence: 1 MRALSFLMKLRNKRKTKMR.....FCHASKENFTPLKEGNPNC 631

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3104	100.0	631	ABG60843	Mouse receptor for
2	3104	100.0	794	AAU11437	Mouse hyaluronic a
3	3098	99.8	631	AAU11437	RHAMM 1-2a isoform
4	3086.5	99.4	630	AAW39166	Mouse RHAMM protei
5	2952.5	95.1	606	AAW39167	Receptor for hyalu
6	2374.5	76.5	713	AAU11438	Rat hyaluronic aci
7	2295.5	74.0	475	AAU11438	Hyaluronan recepto
8	1953	62.9	725	AAW39165	Human RHAMM protei
9	1953	62.9	725	ABG60842	Human receptor for
10	1952	62.9	725	AAU11436	Human hyaluronic a

11	1205	38.8	351	17	AAW01052	Human umbilical ve
12	421.5	13.6	1948	22	ABG21233	Novel human diagno
13	414	13.3	1411	17	ABW02258	Nucleolar/endosoma
14	396.5	12.8	1690	22	ABB61144	Drosophila melanog
15	396.5	12.8	1690	22	ABB61173	Drosophila melanog
16	391	12.6	2954	20	AAW01632	Anino acid sequenc
17	388.5	12.5	1939	23	ABB77096	Human alpha-myosin
18	387	12.5	1886	19	AAW54241	Rattus norvegicus
19	386	12.4	1392	20	AAW06999	Restin protein seq
20	384	12.4	1960	22	AAW78854	Human protein seq
21	384	12.4	1979	21	AAW18171	Plasmodium falcipa
22	384	12.4	2143	22	ABG01716	Novel human diagno
23	383	12.3	1427	12	AAW01534	Human 160kD mediat
24	380.5	12.3	1963	22	AAW79838	Human protein seq
25	375.5	12.1	1988	22	AAW40999	Human polypeptide
26	375.5	12.1	1988	22	AAW41000	Human polypeptide
27	375	12.1	1388	23	AAU79590	Human kinesin moto
28	375	12.1	1388	23	AAE14400	Human kinesin supe
29	374	12.0	1489	22	ABB59948	Drosophila melanog
30	370.5	11.9	931	22	AAW79504	Human protein seq
31	370.5	11.9	990	22	AAW78520	Human protein seq
32	363	11.7	1857	23	AAU84350	Protein MYH11 diff
33	362	11.7	795	23	ABB77430	Human tumour marke
34	362	11.7	885	16	AAW66930	AMML chromosome in
35	362	11.7	976	22	AAG66581	Human SCP-1 mutein
36	362	11.7	2056	22	ABB59344	Drosophila melanog
37	360.5	11.6	1354	18	AAW23654	Physiologically ac
38	360.5	11.6	1354	19	AAW71020	A modified rho tar
39	360.5	11.6	1354	20	AAW07082	Renal cancer assoc
40	360	11.6	2383	23	ABG65631	Human breast speci
41	358	11.5	2117	22	AAU32040	Novel human secret
42	358	11.5	2192	18	AAW21732	LexA/NuMA fusion p
43	358	11.5	2272	18	AAW21731	GAL4/HA/NuMA fusio
44	358	11.5	3248	17	AAW99795	Kinetochore protei
45	357.5	11.5	2101	15	AAW47173	Sequence of the in

ALIGNMENTS

RESULT 1

ABG60843

ID ABG60843 standard; Protein; 631 AA.

XX

AC ABG60843;

XX

DT 13-AUG-2002 (first entry)

XX

DE Mouse receptor for hyalauroan-mediated motility (RHAMM).

XX

KW Tissue disorder; response-to-injury process; cell proliferating;

KW hyaluronic acid; HA; receptor for hyalauroan-mediated motility;

KW RHAMM; inflammatory neurological disorder; Parkinson's disease;

KW Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;

KW inflammatory dermatosis; psoriasis; inflammatory bowel disease;

KW stenosis; restenosis; cancer; kidney fibrosis; inflammatory lung disease;

KW emphysema; asthma; cystic fibrosis; obesity; obesity related disease;

KW lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;

KW tissue transplantation; stroke; inflammatory response; fibrotic response;

KW medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;

KW myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;

KW septic shock; thyroiditis; retinopathy.

XX

OS Mus musculus.

XX

PN WO200228415-A1.

XX

PD 11-APR-2002.

XX

PF 05-OCT-2000; 2000WO-IB01534.

XX

PR 05-OCT-2000; 2000WO-IB01534.

XX

PA (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.

PI Turley EA, Cruz TF;

XX WPI: 2002-435298/46.

PT Treating tissue disorder associated with response-to-injury process or  
PT proliferating cells in mammals, e.g. fibrosis, inflammation, by  
PT administering a compound that alters activity of transition molecules  
PT within a cell

XX Disclosure: Fig 50; 215pp; English.

PS The invention describes a method of treating a tissue disorder associated  
XX with response-to-injury process or proliferating cells in a patient,  
XX comprising administering a polypeptide (I) which binds hyaluronic acid  
XX (HA), an antibody which binds one of domains DI-D5 of Receptor for  
XX hyaluronan-mediated motility (RHAMM), a polypeptide fragment encoding  
XX any of DI-D5 of RHAMM, or a vector which expresses antisense RHAMM,  
XX antibodies or a polypeptide fragment. The method is useful for treating a  
XX patient with an inflammatory neurological disorder such as Parkinson's  
XX disease, Alzheimer's disease, arthritis including rheumatoid arthritis,  
XX osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),  
XX inflammatory bowel disease, stenosis or restenosis, cancer, kidney  
XX fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic  
XX disease), obesity or obesity related diseases, lupus, cardiovascular  
XX disease (e.g. atherosclerosis), and wound especially surgical excision  
XX adhesions, to prevent scar and also for treating or preventing diabetes  
XX mellitus. The method is also useful for treating tissue transplantation  
XX (e.g. skin grafts), stroke, inflammatory responses or fibrotic response  
XX associated with medical implants such as hip implants, vascular wraps and  
XX catheters), inflammatory diseases such as AIDS, myocardial and hepatic  
XX fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,  
XX hepatitis, septic shock, thyroiditis, and retinopathy. This sequence  
XX represents a receptor for hyaluronan-mediated motility protein used in  
XX the method of treating a tissue disorder described in the invention.

SQ Sequence 631 AA;

Query Match 100.0%; Score 3104; DB 23; Length 631;

Best Local Similarity 100.0%; Pred. No. 6.1e-153;

Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALSLEMLKRNKRETKMRSMVMVQEGMELKQATQDLSKQVLEKLSIEKE 60

DB 1 MRALSLEMLKRNKRETKMRSMVMVQEGMELKQATQDLSKQVLEKLSIEKE 60

QY 61 KIDKCEKTEKILEYQIEISCASDVQCKVDIAQLEEDLKEKDRILSLKQLEENITFS 120

DB 61 KIDKCEKTEKILEYQIEISCASDVQCKVDIAQLEEDLKEKDRILSLKQLEENITFS 120

QY 121 KQIEDLVKCOLLETERDNLVSKDRERATLSAEMQILTERLALEROEYKLOQKELQSO 180

DB 121 KQIEDLVKCOLLETERDNLVSKDRERATLSAEMQILTERLALEROEYKLOQKELQSO 180

QY 181 SLLOQEKELSLARLQOLQCSFQEMTSEKRVFKEELKLAELDAVQKQEEQSERLVKQLE 240

DB 181 SLLOQEKELSLARLQOLQCSFQEMTSEKRVFKEELKLAELDAVQKQEEQSERLVKQLE 240

QY 241 EERKSTAEQUTRLDNLRLKEVELEKHTAAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQ 300

DB 241 EERKSTAEQUTRLDNLRLKEVELEKHTAAHAQAIIIAQEKYNDTAQSLRDVTAQLESVQ 300

QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360

DB 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360

QY 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQLESQEKYNDTAQLESQEKYNDTAQLESQ 420

DB 361 SVQEKYNDTAQSLRDVTAQLESQEKYNDTAQLESQEKYNDTAQLESQEKYNDTAQLESQ 420

QY 421 TAESTNQEYARMVQDLQNRSTLKEEIKETISFLEKITDNLQNRQOQDEDFRKQLEEK 480

DB 421 TAESTNQEYARMVQDLQNRSTLKEEIKETISFLEKITDNLQNRQOQDEDFRKQLEEK 480

DB 421 TAESTNQEYARMVQDLQNRSTLKEEIKETISFLEKITDNLQNRQOQDEDFRKQLEEK 480

QY 481 KRTAEKENVMTELTMEINKWRLLYEELYEKTPTFPQOQLDAFAEAKOALLNEHGATOEQLN 540

DB 481 KRTAEKENVMTELTMEINKWRLLYEELYEKTPTFPQOQLDAFAEAKOALLNEHGATOEQLN 540

QY 541 KTRDSYAQLLGHONLKQKTKHVVKLKDENSQKSEVSKLRSQKQNLRLQSGELDKA 600

DB 541 KTRDSYAQLLGHONLKQKTKHVVKLKDENSQKSEVSKLRSQKQNLRLQSGELDKA 600

QY 601 LGIRHFDPSKAFCHASKENFTPLKSGNPCC 631

DB 601 LGIRHFDPSKAFCHASKENFTPLKSGNPCC 631

RESULT 2

AAU11437

ID AAU11437 standard; Protein; 794 AA.

AC AAU11437;

DT 12-MAR-2002 (first entry)

XX Mouse hyaluronic acid binding protein RHAMM.

DE Mouse; hyaluronic acid binding protein; RHAMM; gene therapy;

KW receptor for HA mediated mobility; immunosuppressive; cytostatic.

KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;

OS Mus musculus.

XX Key Location/Qualifiers

FT Misc-difference 55 /note= "Encoded by AAC"

FT Misc-difference 71 /note= "Encoded by ACG"

FT Misc-difference 89 /note= "Encoded by CAA"

FT Misc-difference 91 /note= "Encoded by CAC"

FT Misc-difference 540 /note= "Encoded by ACT"

FT Misc-difference 668 /note= "Encoded by GAT"

PN WO200180899-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-CA00533.

XX 20-APR-2000; 2000US-198613P.

XX (CANG-) CANGENE CORP.

PI Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;

XX WPI: 2002-075094/10.

XX N-PSDB; AAS17497.

XX Protein conjugates that selectively target certain tissues and organs

XX useful for treating and preventing various diseases, comprises

XX glucose-aminoglycan-targeting domain conjugated to a therapeutic

XX protein

PS Claim 6; Page 116; 121pp; English.

XX The invention relates to a conjugate comprising an hyaluronic acid (HA)

XX -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptidic

XX contiguous with, or coupled to a polypeptide conjugated to a therapeutic

XX agent, and the polynucleotides encoding them. Also included is a method

XX for preparation of the HA-binding protein by inserting a first nucleotide

XX sequence encoding a HA-binding protein directly linked to a second

CC nucleotide sequence encoding a therapeutic protein into a suitable  
CC vector, expressing the vector in an acceptable host, purifying conjugate  
CC molecule from host or expression medium. The composition is useful for  
CC altering in vivo the distribution of a therapeutic agent comprising  
CC administering the composition to the animal where conjugate molecule will  
CC distribute primarily in tissues and organs containing high levels of  
CC endogenous HA and for treating mammal with a disorder where a diseased  
CC tissue of the mammal contains high level of HA e.g. rheumatoid  
CC arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic  
CC dosages required also translates into lower immunogenicity of the  
CC conjugated protein as compared to the native protein. As a result,  
CC conjugates improve patient compliance and reduce direct and indirect  
CC costs associated with the drug substance and its administration.  
CC Conjugates allows for the use, where appropriate, of lower, safer,  
CC dosages as compared to the conventional dosage requirements for the  
CC unconjugated corresponding therapeutic agent. Conjugate molecules have an  
CC increased half-life and potency, resulting in prolonged circulation of  
CC the molecule, efficient distribution into the target tissues, and  
CC increased bioavailability. The present sequence represents a RHAMM  
CC protein.

XX Sequence 794 AA;

Query Match 100.0%; Score 3104; DB 23; Length 794;  
Best Local Similarity 100.0%; Pred. No. 7.8e-153;  
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLTQKDLTSGKIVQLGKLVSIKE 60  
Db 164 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLTQKDLTSGKIVQLGKLVSIKE 223  
QY 61 KIDECETKLELEYIQEISCDASQVQCKVDIAQLEEDLKEKDREILSLKQSLSENIITFS 120  
Db 224 KIDECETKLELEYIQEISCDASQVQCKVDIAQLEEDLKEKDREILSLKQSLSENIITFS 283  
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKELQKELQSQ 180  
Db 284 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKELQKELQSQ 343  
QY 181 SLQQEKELSLARLQQLCSFQEMTSKRVFKEELKLAELDAVQOKEEQSERLVKQLE 240  
Db 344 SLQQEKELSLARLQQLCSFQEMTSKRVFKEELKLAELDAVQOKEEQSERLVKQLE 403  
QY 241 PERKSTAQRLTDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300  
Db 404 PERKSTAQRLTDLNLLREKEVELEKHAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 463  
QY 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360  
Db 464 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 523  
QY 361 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLENLTQEKVAMAEKSVEDVQQQIL 420  
Db 524 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLENLTQEKVAMAEKSVEDVQQQIL 583  
QY 421 TAESTNQEYARMVDQLNKSTLKKEIEKITSFLEKITDLKNLRQODEDFRQLEBK 480  
Db 584 TAESTNQEYARMVDQLNKSTLKKEIEKITSFLEKITDLKNLRQODEDFRQLEBK 643  
QY 481 KRTAEKENVMTLTMETINKRLLYBELYEKTKPQQQDLDAFAEAKQALLNEHGATQEQLN 540  
Db 644 KRTAEKENVMTLTMETINKRLLYBELYEKTKPQQQDLDAFAEAKQALLNEHGATQEQLN 703  
QY 541 KIRDSYAQLLGHONLKQIKHVKVLLKDNSQLKSEVSKLRSQLVKRNQELRLQELDKA 600  
Db 704 KIRDSYAQLLGHONLKQIKHVKVLLKDNSQLKSEVSKLRSQLVKRNQELRLQELDKA 763  
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631  
Db 764 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 794

RESULT 3

AAR99675  
ID AAR99675 standard; Protein; 631 AA.  
XX  
AC AAR99675;  
XX  
DT 10-OCT-1996 (first entry)  
XX  
DE RHAMM 1-2a isoform.  
XX  
KW RHAMM 1-2a; receptor for hyaluronic acid mediated motility;  
KW hyaluronan receptor; cell locomotion; cell proliferation;  
KW breast cancer; therapy.

OS Mus sp.

PH Key Location/Qualifiers  
FT Region 55..79  
FT /note= "exon 2A-encoded region"

XX EP721012-A2.

XX 10-JUL-1996.

XX 16-OCT-1995; 95EP-0307310.

PR 14-OCT-1994; 94GB-0020740.

XX (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.  
PA (UYMA-) UNIV MANITOBA.

PI Entwistle J, Turley EA, Zhang S;

DR WPI; 1996-310997/32.

XX N-PSDB; AAT34525.

PT Receptor for hyaluronic acid-mediated motility protein, and DNA  
PT encoding it - useful to treat or prevent diseases associated with  
PT the receptor, e.g. breast cancer

PS Claim 8; Page 50-52; 117pp; English.

CC RHAMM 1-2a (AAR99675) is an alternatively spliced variant of  
CC RHAMM 1 (AAR99673) (receptor for hyaluronic acid mediated  
CC motility), a protein involved in cell locomotion or motility and  
CC cell proliferation and transformation. It differs from RHAMM 1  
CC by an insertion of 25 amino acids (see also AAR99674) between  
CC amino acids 54 and 55 of RHAMM 1, resulting from an alternatively  
CC spliced exon 2A (AAT34502). RHAMM 1-2a is the isoform that is  
CC overexpressed in tumours. Determination of the level of RHAMM  
CC 1-2a in a sample can be used to assess the prognosis of a tumour  
CC (esp. breast cancer) patient. The RHAMM 1-2a protein can also  
CC be used to suppress or control a tumour by modulating the  
CC interaction of cell-associated RHAMM with its ligand.

XX Sequence 631 AA;

Query Match 99.8%; Score 3098; DB 17; Length 631;  
Best Local Similarity 99.7%; Pred. No. 1.3e-152;  
Matches 629; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLTQKDLTSGKIVQLGKLVSIKE 60  
Db 1 MRALSLEMLKLNKRETKMRSMVMVQEGMELKLTQKDLTSGKIVQLGKLVSIKE 60  
QY 61 KIDECETKLELEYIQEISCDASQVQCKVDIAQLEEDLKEKDREILSLKQSLSENIITFS 120  
Db 61 KIDECETKLELEYIQEISCDASQVQCKVDIAQLEEDLKEKDREILSLKQSLSENIITFS 120  
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKELQKELQSQ 180  
Db 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALERQYKELQKELQSQ 180  
QY 181 SLQQEKELSLARLQQLCSFQEMTSKRVFKEELKLAELDAVQOKEEQSERLVKQLE 240

|||||  
Db 181 SLLOQEKELSLRLOQLCSFQEMTSEKNVFKBELKALAEALDAVQOKEEQSERLVKQLE 240  
QY 241 EERKSTAEQLTRDLNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVDTAQLSVQ 300  
Db 241 EERKSTAEQLTRDLNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVDTAQLSVQ 300  
QY 301 EKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQL 360  
Db 301 EKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQL 360  
QY 361 SVOEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQL 420  
Db 361 SVOEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQL 420  
QY 421 TAESTNOEYARMVQDLQNRSTLKEEIKETSSFKETITDLKNLRQODEDFRKOLEEG 480  
Db 421 TAESTNOEYARMVQDLQNRSTLKEEIKETSSFKETITDLKNLRQODEDFRKOLEEG 480  
QY 481 KRTAEKENVMTELTMEINKWRLLYEELYEKTPFQOQLDAFAEAKQALLNEHGATQEQLN 540  
Db 481 KRTAEKENVMTELTMEINKWRLLYEELYEKTPFQOQLDAFAEAKQALLNEHGATQEQLN 540  
QY 541 KIRDSYAQLLGHQNLKQKIKHVVKLDENSQKSEVSKLRSQVSRKQNELRLQGELOKA 600  
Db 541 KIRDSYAQLLGHQNLKQKIKHVVKLDENSQKSEVSKLRSQVSRKQNELRLQGELOKA 600  
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGNPNCC 631  
Db 601 LGIRHFDPSKAFCHASKENFTPLKEGNPNCC 631

RESULT 4  
AAW39166  
ID AAW39166 standard; Protein; 630 AA.  
XX  
AC AAW39166;  
XX  
DT 27-APR-1998 (first entry)  
XX  
DE Mouse RHAMM protein.  
XX  
KW Hyaluronan receptor; receptor for hyaluronic acid mediated motility;  
KW RHAMM; glycosaminoglycan; binding domain; mouse; oncogene; treatment;  
KW growth factor; cell locomotion disorder; dementia; detection;  
KW inflammatory disorder; autoimmune disease; diagnosis; prognosis.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Region 279..382  
FT Binding-site /note= "repeat region"  
FT Binding-site 555..565  
FT Binding-site /note= "Hyaluronic acid binding domain"  
FT Binding-site 577..586  
FT Binding-site /note= "Hyaluronic acid binding domain"  
XX  
PN WO9738098-A1.  
PD  
PD 16-OCT-1997.  
XX  
PF 10-APR-1997; 97WO-CA00240.  
XX  
PR 10-APR-1996; 96GB-0007441.  
XX  
XX (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.  
PA (UTMA-) UNIV MANITOBA.  
XX  
XX Entwistle J, Turley EA;  
XX WPI; 1997-512715/47.  
DR N-PSDB; AAW02801.  
XX

Isolated human receptor for hyaluronic acid mediated motility - used  
to develop products for treating e.g. tumours, inflammatory  
disorders, dementia, AIDS, diabetes and auto-immune diseases  
Dislosure; Page 46; 66pp; English.  
XX  
This sequence represents the mouse hyaluronan receptor which is also  
known as the receptor for hyaluronic acid mediated motility (RHAMM).  
Hyaluronan is a large glycosaminoglycan that is ubiquitous in the  
extracellular matrix and whose synthesis has been linked to cell  
migration, growth and transformation. It interacts with cell surfaces via  
specific protein receptors, e.g. RHAMM, that mediate many biological  
effects. The RHAMM/hyaluronic acid interaction is involved in  
oncogene- and growth factor-mediated cell locomotion. The products can be  
used in the treatment of disorders involving cell locomotion, e.g. tumour  
invasion, birth defects, acute and chronic inflammatory disorders,  
Alzheimer's and other forms of dementia, autoimmune diseases, corneal  
Huntington's diseases, AIDS, diabetes, autoimmune diseases, and  
dysplasia and hypertrophies, burns, surgical incisions and adhesions,  
strokes and multiple sclerosis. They can also be used in e.g. CNS and  
spinal cord regeneration, contraception and in vitro fertilisation and  
embryo development. The products can also be used in detection, diagnosis  
and prognosis.  
XX  
SQ Sequence 630 AA;  
Query Match 99.4%; Score 3086.5; DB 18; Length 630;  
Best Local Similarity 99.7%; Pred No. 4.9e-152;  
Matches 629; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
QY 1 MRALSLELMKLRNKRRETKMRSMVMVQOEGMELKLOATQKDLTESKGIYQLEKLSIEKE 60  
Db 1 MRALSLELMKLRNKRRETKMRSMVMVQOEGMELKLOATQKDLTESKGIYQLEKLSIEKE 60  
QY 61 KIDECETEKLLLEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSKQSLEENITPS 120  
Db 61 KIDECETEKLLLEYIQEISCASDQVEKCKVDIAQLEEDLKEKDREILSKQSLEENITPS 120  
QY 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKLOQELQSQ 180  
Db 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKLOQELQSQ 180  
QY 181 SLLOQEKELSLRLOQLCSFQEMTSEKNVFKBELKALAEALDAVQOKEEQSERLVKQLE 240  
Db 181 SLLOQEKELSLRLOQLCSFQEMTSEKNVFKBELKALAEALDAVQOKEEQSERLVKQLE 240  
QY 241 EERKSTAEQLTRDLNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVDTAQLSVQ 300  
Db 241 EERKSTAEQLTRDLNLLREKEVELEKHIAHAQAIIIAQEKYNDTAQSLRDVDTAQLSVQ 300  
QY 301 EKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQL 360  
Db 301 EKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQL 360  
QY 361 SVOEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQL 420  
Db 361 S-VEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQLSESEQEKYNDTAQSLRDVDTAQL 419  
QY 421 TAESTNOEYARMVQDLQNRSTLKEEIKETSSFKETITDLKNLRQODEDFRKOLEEG 480  
Db 420 TAESTNOEYARMVQDLQNRSTLKEEIKETSSFKETITDLKNLRQODEDFRKOLEEG 479  
QY 481 KRTAEKENVMTELTMEINKWRLLYEELYEKTPFQOQLDAFAEAKQALLNEHGATQEQLN 540  
Db 480 KRTAEKENVMTELTMEINKWRLLYEELYEKTPFQOQLDAFAEAKQALLNEHGATQEQLN 539  
QY 541 KIRDSYAQLLGHQNLKQKIKHVVKLDENSQKSEVSKLRSQVSRKQNELRLQGELOKA 600  
Db 540 KIRDSYAQLLGHQNLKQKIKHVVKLDENSQKSEVSKLRSQVSRKQNELRLQGELOKA 599  
QY 601 LGIRHFDPSKAFCHASKENFTPLKEGNPNCC 631  
Db 600 LGIRHFDPSKAFCHASKENFTPLKEGNPNCC 630

CC found, RHAMM 1A (see also AAT34523) and RHAMM 1B (AAT34524), that had  
CC identical translated portions. Increased expression of RHAMM 1  
CC protein is indicative of a poor prognosis for breast cancer. The  
CC protein can be used to suppress or control a tumour by modulating  
CC the interaction of cell-associated RHAMM with its ligand.

XX AAR99673 standard; Protein; 606 AA.

XX AAR99673;

XX 10-OCT-1996 (first entry)

XX Receptor for hyaluronic acid mediated motility RHAMM 1.

KW RHAMM 1; receptor for hyaluronic acid mediated motility;  
KW hyaluronan receptor; cell locomotion; cell proliferation;  
KW breast cancer; therapy.

OS Mus sp.

XX Key Location/Qualifiers

FT Modified-site 91..93  
FT /label= N-glycosylation\_site  
FT Modified-site 258..260  
FT /label= N-glycosylation\_site  
FT Modified-site 279..281  
FT /label= N-glycosylation\_site  
FT Modified-site 300..302  
FT /label= N-glycosylation\_site  
FT Modified-site 321..323  
FT /label= N-glycosylation\_site  
FT Modified-site 342..344  
FT /label= N-glycosylation\_site  
FT Modified-site 373..375  
FT /label= N-glycosylation\_site  
FT Modified-site 413..415  
FT /label= N-glycosylation\_site  
FT Domain 532..542  
FT /label= Hyaluronan\_binding\_domain-I  
FT Domain 553..562  
FT /label= Hyaluronan\_binding\_domain-I  
FT Modified-site 594..596  
FT /label= N-glycosylation\_site

XX EP721012-A2.

XX 10-JUL-1996.

XX 16-OCT-1995; 95EP-0307310.

XX 14-OCT-1994; 94GB-0020740.

XX (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.  
XX (UYMA-) UNIV MANITOBA.

XX Entwistle J, Turley EA, Zhang S;

XX WPI; 1996-310997/32.

XX N-PSDB; AAT34499;

XX N-PSDB; AAT34500;

XX N-PSDB; AAT23423;

XX N-PSDB; AAT34524.

XX Receptor for hyaluronic acid-mediated motility protein, and DNA

PT encoding it - useful to treat or prevent diseases associated with  
PT the receptor, e.g. breast cancer  
XX  
PS Claim 2; Page 40-42; 117pp; English.  
XX  
CC RHAMM 1 (AAR99673), or Receptor for Hyaluronic Acid Mediated Motility,  
CC is a hyaluronan receptor protein which is involved in cell locomotion  
CC or motility and cell proliferation and transformation. Its amino  
CC acid sequence was deduced from a cDNA clone (AAT34499) obt'd. from  
CC murine 3T3 cells and a genomic sequence (AAT34500) from a mouse  
CC fibroblast genomic library. 2 Alternative mRNAs for RHAMM 1 were

XX SQ Sequence 606 AA;

Query Match 95.1%; Score 2952.5; DB 17; Length 606;  
Best Local Similarity 95.7%; Pred. No. 4e-145;  
Matches 604; Conservative 2; Mismatches 0; Indels 25; Gaps 1;

Qy 1 MRALSLELMKLRNKRKTRMRSMMVKQEGMELKQATQKDLTSGKGIQVLEKGLVSIKIE 60  
Db 1 MRALSLELMKLRNKRKTRMRSMMVKQEGMELKQATQKDLTSGKGIQVLEKGLVSIKIE 54  
Qy 61 KIDKCEKLEKLEYIOEISCASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENITFS 120  
Db 55 -----CASDQVEKCKVDIAQLEEDLKEKDREILSLKQSLSEENITFS 95  
Qy 121 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEREQVEYKLOQKELQSO 180  
Db 96 KOIEDLTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEREQVEYKLOQKELQSO 155  
Qy 181 SILQOEKELSLARLQOOLCSFQEMTSKRVKKEELKALAEALDAVQOKEEQSERLVKQLE 240  
Db 156 SILQOEKELSLARLQOOLCSFQEMTSKRVKKEELKALAEALDAVQOKEEQSERLVKQLE 215  
Qy 241 EERKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 300  
Db 216 EERKSTAEQLTRLDNLLREKEVELEKHIAHAQAAILIAQEKYNDTAQSLRDVTAQLESVQ 275  
Qy 301 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 360  
Db 276 EKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLE 335  
Qy 361 SVQEKYNDTAQSLRDVSAQLESYKSTLKEIEDLKLENLTLOEKVAMAEKSVEDVQOQIL 420  
Db 336 SVQEKYNDTAQSLRDVTAQLESYKSTLKEIEDLKLENLTLOEKVAMAEKSVEDVQOQIL 395  
Qy 421 TAESTNQEYARMVQDLQNRSTLKEEIEKITSFLEKTDLNQNRQOQDEDFRKQLEEK 480  
Db 396 TAESTNQEYARMVQDLQNRSTLKEEIEKITSFLEKTDLNQNRQOQDEDFRKQLEEK 455  
Qy 481 KRTAEKNVMTLWELINKRWLLYBELYEKTPFOQDLDAFAEAKOALLNEHGATQEQLN 540  
Db 456 KRTAEKNVMTLWELINKRWLLYBELYEKTPFOQDLDAFAEAKOALLNEHGATQEQLN 515  
Qy 541 KIRDSYAQLLGHQNLKQIKHVVKLKDENSQKLSVSKLSRSLQVVRKQNELRLQGLDKA 600  
Db 516 KIRDSYAQLLGHQNLKQIKHVVKLKDENSQKLSVSKLSRSLQVVRKQNELRLQGLDKA 575  
Qy 601 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 631  
Db 576 LGIRHFDPSKAFCHASKENFTPLKEGPNCC 606

RESULT 6

AAU11438

ID AAU11438 standard; Protein; 713 AA.

XX AAU11438;

XX 12-MAR-2002 (first entry)

XX Rat hyaluronic acid binding protein RHAMM.

KW Rat; hyaluronic acid binding protein; RHAMM; gene therapy;  
KW receptor for HA mediated motility; immunosuppressive; cytostatic.  
KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;  
OS Rattus norvegicus.  
XX







Db 458 ----- 457  
QY 360 ESVOEKYNDTAQSLRDVSAQLESYKSTLKEETEDLKLNTLQEKVAMAEKSVEDVQOOI 419  
Db 458 -----AQFESYKALTASEIEDLKLENSLQEKAAKAGNAEDVQHOI 499  
QY 420 LTAESTNOEYARMVODLQNRSTLKEEIEKITSSFLKTTDLKNLQROQDEDFRQLEEK 479  
Db 500 LATESSENOEYVRLMLDLQTKSALKETEITVSLFKITDLDLQNLKQOEEDFRQLEDE 559  
QY 480 GKRTAKENVMPELMEINKWRLLYEELYEKTPFOQQLDAFAEKQALLNEHGATQEOI 539  
Db 560 EGKRAEKENTTAELTEINKWRLLYEELYNKTPFOLQDAFEVEKQALLNEHGAQEOI 619  
QY 540 NKIRDSYAKLLGHONLQKIKHVVKLDKENSOLKSVSKRLQKQAKKQSEYKQLEELNK 599  
Db 620 NKIRDSYAKLLGHONLQKIKHVVKLDKENSOLKSVSKRLQKQAKKQSEYKQLEELNK 679  
QY 600 ALGIRHFDPSKAFCHASKENF---TPLKEGPNPC 630  
Db 680 VLGIKHFDPSPKAFCHASKENFALKTPKKEGNTNC 713  
  
RESULT 9  
ID ABG60842  
XX ABG60842 standard; Protein; 725 AA.  
XX AC  
XX ABG60842;  
DT 13-AUG-2002 (first entry)  
XX Human receptor for hyaluronan-mediated motility (RHAMM).  
XX  
XX Tissue disorder; response-to-injury process; cell proliferating;  
KW hyaluronic acid; HA; receptor for hyaluronan-mediated motility;  
KW RHAMM; inflammatory neurological disorder; Parkinson's disease;  
KW Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;  
KW inflammatory dermatosis; psoriasis; inflammatory bowel disease;  
KW stenosis; restenosis; cancer; kidney fibrosis; inflammatory lung disease;  
KW emphysema; asthma; cystic fibrosis; obesity; obesity related disease;  
KW lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;  
KW tissue transplantation; stroke; inflammatory response; fibrotic response;  
KW medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;  
KW myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;  
KW septic shock; thyroiditis; retinopathy.  
XX  
OS Homo sapiens.  
XX  
XX WO200228415-A1.  
XX  
XX 11-APR-2002.  
XX  
XX 05-OCT-2000; 2000WO-IB01534.  
XX  
XX 05-OCT-2000; 2000WO-IB01534.  
XX  
XX (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.  
XX  
XX Turley EA, Cruz TF;  
XX  
XX WPI; 2002-435298/46.  
XX  
XX Treating tissue disorder associated with response-to-injury process or  
XX proliferating cells in mammals, e.g. fibrosis, inflammation, by  
XX administering a compound that alters activity of transition molecules  
XX within a cell  
XX  
XX Example 30; Fig 50; 215pp; English.  
XX  
XX The invention describes a method of treating a tissue disorder associated  
XX with response-to-injury process or proliferating cells in a patient,  
XX comprising administering a polypeptide (I) which binds hyaluronic acid

CC (HA), an antibody which binds one of domains DI-D5 of Receptor for  
CC hyaluronan-mediated motility (RHAMM), a polypeptide fragment encoding  
CC any of DI-D5 of RHAMM, or a vector which expresses antisense RHAMM,  
CC antibodies or a polypeptide fragment. The method is useful for treating a  
CC patient with an inflammatory neurological disorder such as Parkinson's  
CC disease, Alzheimer's disease, arthritis including rheumatoid arthritis,  
CC osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),  
CC inflammatory bowel disease, stenosis or restenosis, cancer, kidney  
CC fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic  
CC fibrosis), obesity or obesity related diseases, lupus, cardiovascular  
CC disease (e.g. atherosclerosis), and wound especially surgical excision  
CC adhesions, to prevent scar and also for treating or preventing diabetes  
CC mellitus. The method is also useful for treating tissue transplantation  
CC (e.g. skin grafts), stroke, inflammatory responses or fibrotic response  
CC associated with medical implants such as hip implants, vascular wraps and  
CC catheters), inflammatory diseases such as AIDS, myocardial and hepatic  
CC fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,  
CC hepatitis, septic shock, thyroiditis, and retinopathy. This sequence  
CC represents a receptor for hyaluronan-mediated motility protein used in  
CC the method of treating a tissue disorder described in the invention.  
XX  
XX Sequence 725 AA;  
SQ  
  
Query Match 62.9%; Score 1953; DB 23; Length 725;  
Best Local Similarity 64.8%; Pred. No. 2.2e-93;  
Matches 411; Conservative 49; Mismatches 86; Indels 88; Gaps 3;  
  
QY 1 MRALSLELMKLNKRETKMRSMVMKQEGMELQATOKDLSKGVQLECKLVSIKIE 60  
Db 164 LRLSLELMKLNKRETKMRSMVMKQEGMELQATOKDLSKGVQLECKLVSIKIE 223  
QY 61 KIDKCEKTEKLEIYQIEISCASQDQVEKCKVDIALEEDLKEKDREITLSLKQSEENIT-F 119  
Db 224 KIDKCEKTEKLEIYQIEISCASQDQVEKCKVDIALEEDLKEKDREITLSLKQSEENITVL 283  
QY 120 SKQIEDLVKCOLLETERDNLVSKDRAETLSAEMQILTERLALERQYKQLEKQLOS 179  
Db 284 SKQVEDLVKCOLLEKEDHVNRRNHNENIENAEQNLKQFLEQOQSHKQOQLEQI 343  
QY 180 OSLLQOQKELSLARQOQSCFOEEMTSEKNVFEELKALAEADAVQOQSEKSERLVKOL 239  
Db 344 DLSLQOQKELSLARQOQSCFOEEMTSEKNVFEELKALAEADAVQOQSEKSERLVKOL 403  
QY 240 EEEKSTAEQTLRLNLLREKEVELEKHAHAQAAILTAQEKYNDTAQSLRDVTAQLESV 299  
Db 404 EEEKSTAEQTLRLNLLREKEVELEKHAHAQAAILTAQEKYNDTAQSLRDVTAQLESV 457  
QY 300 QEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQLESQEKYNDTAQSLRDVTAQ 359  
Db 458 ----- 457  
QY 360 ESVOEKYNDTAQSLRDVSAQLESYKSTLKEETEDLKLNTLQEKVAMAEKSVEDVQOOI 419  
Db 458 -----AQFESYKALTASEIEDLKLENSLQEKAAKAGNAEDVQHOI 499  
QY 420 LTAESTNOEYARMVODLQNRSTLKEEIEKITSSFLKTTDLKNLQROQDEDFRQLEEK 479  
Db 500 LATESSENOEYVRLMLDLQTKSALKETEITVSLFKITDLDLQNLKQOEEDFRQLEDE 559  
QY 480 GKRTAKENVMPELMEINKWRLLYEELYEKTPFOQQLDAFAEKQALLNEHGATQEOI 539  
Db 560 EGKRAEKENTTAELTEINKWRLLYEELYNKTPFOLQDAFEVEKQALLNEHGAQEOI 619  
QY 540 NKIRDSYAKLLGHONLQKIKHVVKLDKENSOLKSVSKRLQKQAKKQSEYKQLEELNK 599  
Db 620 NKIRDSYAKLLGHONLQKIKHVVKLDKENSOLKSVSKRLQKQAKKQSEYKQLEELNK 679  
QY 600 ALGIRHFDPSKAFCHASKENF---TPLKEGPNPC 630  
Db 680 VLGIKHFDPSPKAFCHASKENFALKTPKKEGNTNC 713  
  
RESULT 10

AAU11436	AAU11436 standard; Protein; 725 AA.
XX	AAU11436;
XX	AAU11436;
XX	12-MAR-2002 (first entry)
XX	Human hyaluronic acid binding protein RHAMM.
DE	Human; hyaluronic acid binding protein; RHAMM; gene therapy;
XX	receptor for HA mediated mobility; immunosuppressive; cytostatic.
KW	conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;
KW	
KW	
XX	Homo sapiens.
OS	
XX	Key
XX	Location/Qualifiers
XX	Misc-difference 278
FT	/note= "Encoded by GAA"
FT	
FT	Misc-difference 299
FT	/note= "Encoded by AAA"
FT	
FT	Misc-difference 323
FT	/note= "Encoded by AAA"
FT	
FT	Misc-difference 331
FT	/note= "Encoded by CAG"
FT	
XX	
XX	WO200180899-A2.
PN	
XX	01-NOV-2001.
XX	
XX	20-APR-2001; 2001WO-CA00533.
PD	
XX	
XX	20-APR-2000; 2000US-198613P.
PR	
XX	(CANG-) CANGENE CORP.
XX	
XX	Woloski BMP, Williams AM, Sereda TJ, Wiebe DJ;
PI	
XX	WPI; 2002-075094/10.
DR	N-PSDB; AAS17496.
DR	
XX	
XX	Protein conjugates that selectively target certain tissues and organs
PT	useful for treating and preventing various diseases, comprises
PT	glucose-aminoglycan-targeting domain conjugated to a therapeutic
PT	protein
XX	
XX	Claim 6; Page 115; 121pp; English.
PS	
PS	
XX	
CC	The invention relates to a conjugate comprising an hyaluronic acid (HA)
CC	-binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide
CC	contiguous with, or coupled to a polypeptide conjugated to a therapeutic
CC	agent, and the polynucleotides encoding them. Also included is a method
CC	for preparation of the HA-binding protein by inserting a first nucleotide
CC	sequence encoding a HA-binding protein directly linked to a second
CC	nucleotide sequence encoding a therapeutic protein into a suitable
CC	vector, expressing the vector in an acceptable host, purifying conjugate
CC	molecule from host or expression medium. The composition is useful for
CC	altering in vivo the distribution of a therapeutic agent comprising
CC	administering the composition to the animal where conjugate molecule will
CC	distribute primarily in tissues and organs containing high levels of
CC	endogenous HA and for treating mammal with a disorder where a diseased
CC	tissue of the mammal contains high level of HA e.g. rheumatoid
CC	arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic
CC	dosages required also translates into lower immunogenicity of the
CC	conjugated protein as compared to the native protein. As a result,
CC	conjugates improve patient compliance and reduce direct and indirect
CC	costs associated with the drug substance and its administration.
CC	Conjugates allows for the use, where appropriate, of lower, safer,
CC	dosages as compared to the conventional dosage requirements for the
CC	unconjugated corresponding therapeutic agent. Conjugate molecules have an
CC	increased half-life and potency, resulting in prolonged circulation of
CC	the molecule, efficient distribution into the target tissues, and
CC	increased bioavailability. The present sequence represents a RHAMM
CC	protein.
CC	

XX	Sequence	725 AA;
50	Query Match	62.9%; Score 1952; DB 23; Length 725;
	Best Local Similarity	64.7%; Pred. No. 2.5e-93;
	Matches 410; Conservative	51; Mismatches 85; Indels 88; Gaps
QY	1	MRALSLEMLRNKRKTKMRSMVMYQGBMELKLAQTKOKLJTESKGIQVLEGKLVISIEKE 60
DB	164	LRILSELMLRNKRKTKMRGMMAKQGBMELKLVQVTRSEESQGIQAQLEGKLVISIEKE 122
QY	61	KIDKCEKTEKLLVYQIBISCASDOVEKCKVDIAQLEEDLKEKQREILSLKQSLLEENIT-F 119
DB	224	KIDKSEKTEKLLVYIEISCASDOVEKCKVDIAQLEENLKEKNDILSLKQSLLEENIT-F 119
QY	120	SKQEDTVKCOLLETERDNLVSKDRERAETLSAEMQILTERLALEROEYKLOQRELOS 179
DB	284	SKQVEDLVKCOLLETEREKEDVHVRNRHNNLNAMQNLQKQFVLEQREHEKLOQRELOI 343
QY	180	QSLLOQKEKLSARLQOQLCSFQEBMTSEKVNFKELKALAEALDAVQOKEQSERLVKQL 239
DB	344	DSLLOQKEKLSLSSHLQKLCFSQEBMVKENKLFEEELKQTLDELKDKLOQKEQEAERLVKQL 403
QY	240	EEERKSTAEQULTRLDNLLRKEVELEKHHAAHAQAALIAQEKYNDTAQSLRDVTAQLESV 299
DB	404	EEEAASRAEELKLEELKGAELKESAAHTQATLLQEKYDSMVQSLDYT----- 457
QY	300	QEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQLESEQEKYNDTAQSLRDVTAQL 359
DB	458	----- 457
QY	360	ESVOEKYNDTAQSLRDVSAQLESYKSSTLKEIEDLKLLENLTLOEKVAMAEKSYEDVQOOI 419
DB	458	-----AQFESYKALTASEIEDLKLSSSLQEKAAKAGKNAEDVQHOI 499
QY	420	LTAESTNOEYARMVDLQNRSTLKEBEIKETISFLBKITDLKNQLRQOQEDPRKOLEEK 479
DB	500	LATESSQEYVYMLDLQTKSALKETIKETISFLQKITDLQNLQKQOEDPRKOLEDE 559
QY	480	GKRTAEKENYMTLTMENKWRLLYELYEKTKPFQOQLDAFAEAKOALLNEHGATQEQ 539
DB	560	EGRKAEKENTTAELTEINKWRLLYELLYNKTPFQQLDAFEVQKALLNEHGAAQEQ 619
QY	540	NKIRDSYAQLGHQNLKQKIKHVYKLDKNSQLKSEVSKLRSQLVKRRKQNELRLOQELDK 599
DB	620	NKIRDSYAKLGHQNLKQKIKHVYKLDKNSQLKSEVSKLRCQLAKKKQSETKLQELN 679
QY	600	ALGTRHFDPSKAFCHASKENF---TPLKEGNPC 630
DB	680	VLGIKHFDPKAFHESKENFALKTPLKEGNTNC 713
XX	RESULT 11	
XX	AAW01052	
XX	ID	AAW01052 standard; Protein; 351 AA.
XX	AC	AAW01052;
XX	DT	31-JAN-1997 (first entry)
XX	DE	Human umbilical vein epithelium-derived hyaluronan receptor.
XX	KW	Umbilical; diagnosis; detection; hyaluronan receptor; antibody;
XX	KW	cancer; inflammation; angiogenesis; invasive; leukaemia; lymphoma;
XX	KW	proliferation; vascularisation.
XX	OS	Homo sapiens.
XX	PN	W09268549-A2.
XX	PD	19-SEP-1996.
XX	PF	08-MAR-1996; 96WO-USO3193.



Db 1013 TLDDQLMEEDKVTNLTAKTKLEQVDDLEGSLEQEKKLCMDLERAKRKLGLDLKLAQES 1072  
 QY 221 ELDAVQOEEQSERLVKO-----LE---EERKSTAEOLTRLDNLREKEVELEKHAH 271  
 Db 1073 TMDTENDQOLNEKLLKKEFEMSNLOGKIEDEQALAMQOLKIKELQARIIELEEEIEAE 1132  
 QY 272 AQAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESQEKYNDTAQSL 331  
 Db 1133 RASRAKAEKQSDLSRELEISERLE-----EAGGATSAQIELNKKREAEFOKMR 1182  
 QY 332 RDVTAQLESQEKYNDTAQSLRDVTAQLESVQEKYNDTAQSLRDVTAQLESYKSKSTLKEI 391  
 Db 1183 RD-----LEESTLQHEATAAALRKHHA--DSVAE-IGKQIDSLQRYVKQKLEKEKSELKMEI 1235  
 QY 392 EDLKNLTTLQEKVAMAEKSVEDVQOQILTAESTNQEVARMVQDQONRSTLKEEIKKIT 451  
 Db 1236 NDLASNMETVSKAKANPEKMCFTLEDQISEIKTKEEEOORLINELSAOKARLHTEGFE 1295  
 QY 452 -----SSFLEKITDLKNLRQO-----DED-FRKQLE 477  
 Db 1296 RQLEKADAMVSQSRGKQFTQOIEELKRLQEEETKAKSTLAHALQASARHDCDLLREQYE 1355  
 QY 478 EKGKTAENVTMTLWEINKWRLLY-----EELYEKTTPQOQLD-----519  
 Db 1356 EQEAKAELQRMGSKANSEVAQWRTKYETDAIQRTTEELEEAQKLAQRLQDAEHEVAVN 1415  
 QY 520 -----AFEAQKQALLNE-----HGATQEOQLN 540  
 Db 1416 SKASLETKQRLQNEVEDLMIDVERSNACIALDKKORNFEDKVLAEWKQRYEETQAELE 1475  
 QY 541 -----KIRDSYAQLLGH-----ONLKOKI-----KHVVKLKD 567  
 Db 1476 ASQKESRSLSTELFKVKNAYEESLDHLETLKRENKNLQOEISDLTEQIAEGGKHTELEK 1535  
 QY 568 ENSQLKSEVSKLRSOLVRKQNELRLQELGELDKALGIR 604  
 Db 1536 VKKQDHEKSELOSL-----EEAEASLEHEEGKILRIQ 1569

RESULT 13

ID AAW02258 standard; Protein; 1411 AA.  
 AC AAW02258;  
 DT 09-MAR-1997 (first entry)

DE Nucleolar/endosomal auto-antigen p162.

KW Auto-antibody; p162; rheumatic disease; antigen; diagnosis;  
 gene therapy.

OS Homo sapiens.

PN DE19515514-C1.

PD 12-SEP-1996.

PF 27-APR-1995; 95DE-1015514.

PR 27-APR-1995; 95DE-1015514.

PA (PRIV-) PRIVATES INST IMMUNOLOGIE & MOLEKULARGEN.

PI Renz M, Seelig HP;

DR WPI; 1996-403153/41.

XX N-PSDB; AAT58751.

PT DNA encoding nucleolar-endosomal auto-antigen - useful for exact  
 diagnosis of rheumatic disease, in gene therapy and for removal of  
 specific auto-antibodies

PS Claim 1; Fig 2; 15pp; German.  
 XX Transformed cells can be cultured to produce the antigen p162, for use  
 in exact (differential) diagnosis of rheumatic disease, i.e. they  
 can detect, in immunoassays, Western blots, etc., rheumatism-  
 specific auto-antibodies. The antigen can be used therapeutically,  
 in the removal of auto-antibodies from the circulation, or when  
 coupled to a cytotoxin, the elimination of auto-antibody-  
 producing lymphocytes.

XX Sequence 1411 AA;

Query Match 13.3%; Score 414; DB 17; Length 1411;  
 Best Local Similarity 23.0%; Pred. No. 1.5e-13;  
 Matches 168; Conservative 150; Mismatches 220; Indels 192; Gaps 28;

QY 7 ELMKLRN-----KRETKMRSMVMYKQEG--MEUK-----LOAT--OKDLTSEK 44  
 Db 289 ELQKLKSSVNELTQKNQITLNLAKKEQDYTKLEEKHNEESVSKKNIQATLHKQDL---- 344  
 QY 45 GKIVOLEGKLVSIKE-----KIDKCE-TEKLLEYIQEISCASDOVEKCKVDIAQLEE 97  
 Db 345 -DCQQLQSRRLSASETSLRHIVHSELSEKGEATQKLEELSEV---ETKQHLKAEFKQLQ 400  
 QY 98 DLKEDREILSKQSLSEENITFSKOIEDLVKCOLLETEDNLVSKDRERAEFTLSAEMQI 157  
 Db 401 QREKEQHGLQLOSEINQ-----LHKKLETERQLGEAHR-----LKEQRL 443  
 QY 158 LTERLALREQYEKELQQLQELQSLQELQKELSAKLQQL-----CSFQEMTSE 207  
 Db 444 SSEKLMDREQQVADQLKLSLEELQKKEKVTNSTELQHLQDKTKQHQHQQALQOSTTAK 503  
 QY 208 KNVFEELKLALADAVQOKEEQSERLVQ-----LEERK-----STA 247  
 Db 504 LREAQNDLEQVLRQIGDKDKQKQIONLEALLOKSKENISLEKEREEDLYAKIAGEGETAVL 563  
 QY 248 EQLTRLDNLLREKEVELEKHTAAHQAAILIAQEKYNDTAQSLRDVTAQLESVQEKYNDTA 307  
 Db 564 NQLQEKNHTLQEQVTQLTEKLNQSHKQAOENLHDQVQEQK---AHLRAAQDRVLSLE 620  
 QY 308 QSLRDVTAQLESQEKYND-----TAQSLRDVTAQLESQEKYNDT 348  
 Db 621 TSVNELNSQLNESKEKVSQLDIQIKAKTELLLSAEAAKTAQR-ADLQNLHDLTAQNALQDK 679  
 QY 349 AQSLRDVTAQLESVQEKYNDTAQSLRDVSAQLESYKSTL---KEIEDL-----KLENLT 400  
 Db 680 QOELNKITQLDQVTAKLQDKQEHCSQLESHLKYEKYLSEQKTEEGQIKKLEADS 739  
 QY 401 LOEKVAMAEKSVEDVQOQILTAESTNQEVARMVQDQONRSTLKEEIEI---KEITSF--- 454  
 Db 740 LEVK-ASKEQALQDLQOQ-----RQLNTDLELRATELSKQLEMEKEIVSSTRLD 787  
 QY 455 ----LEKITDKNQLRQODEDFRKQLEKGTAKENVMTELTWEINKWRLLYELYEK 510  
 Db 788 LQKKSEALESIKQKLTQKEE--KQILQDFFETLSQET-----KIQHEELNMR 833  
 QY 511 TKPFOQLDAFAEAKQALLNEHGATQEOQLNKIRDSY-----AQLLGHQNL 555  
 Db 834 IQTTVTLEQKVMEKEALMTLSTVTKDLKSVSDSKNSKSEFEKENQKGAAILDLEKT 893  
 QY 556 KQIKHVVKLKDENS-----OLKSEVSKLRSQLVKVR-----KON---589  
 Db 894 CKELKHQLOVQMENTLKEQKELKKSLEKEKEASHQKLELNSMQELOLQAOQNTLKQNEKE 953  
 QY 590 ELRLQGLDK 599  
 Db 954 EQOQLQGNINE 963

RESULT 14  
 ABB61144  
 ID ABB61144 standard; Protein; 1690 AA.  
 XX





**THIS PAGE BLANK (USPTO)**